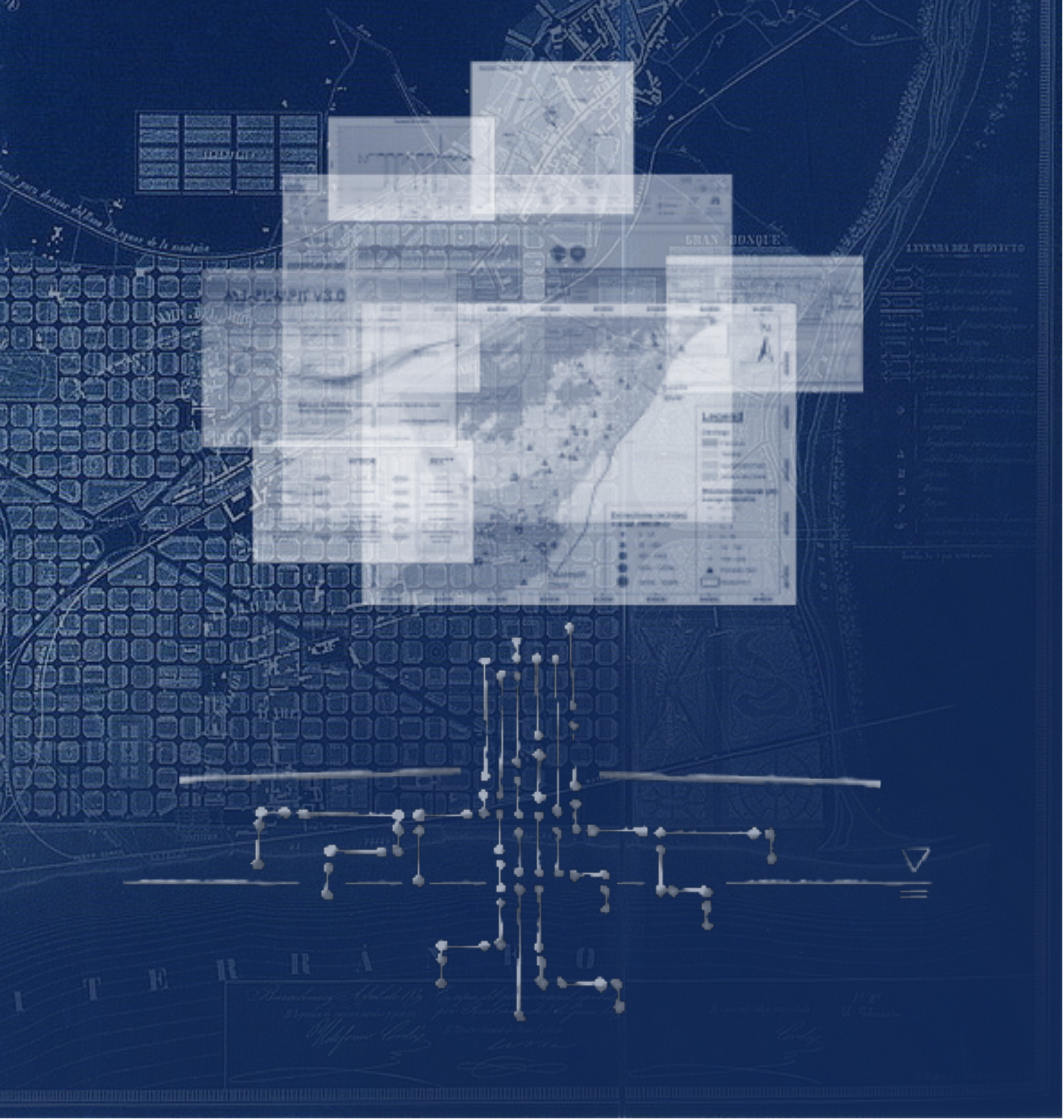
An approach for hydrogeological data management, integration and analysis

Rotman A. Criollo Manjarrez



An approach for hydrogeological data management, integration and analysis

Plataforma de gestió i anàlisi de dades geològiques e hidrogeològiques a la ciutat de Barcelona

Rotman A. Criollo Manjarrez

ADVERTIMENT La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del repositori institucional UPCommons (http://upcommons.upc.edu/tesis) i el repositori cooperatiu TDX (http://www.tdx.cat/) ha estat autoritzada pels titulars dels drets de propietat intel·lectual **únicament per a usos privats** emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei UPCommons o TDX. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a UPCommons (*framing*). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del repositorio institucional UPCommons (http://upcommons.upc.edu/tesis) y el repositorio cooperativo TDR (http://www.tdx.cat/?locale-attribute=es) ha sido autorizada por los titulares de los derechos de propiedad intelectual **únicamente para usos privados enmarcados** en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio UPCommons No se autoriza la presentación de su contenido en una ventana o marco ajeno a UPCommons (*framing)*. Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the institutional repository UPCommons (http://upcommons.upc.edu/tesis) and the cooperative repository TDX (http://www.tdx.cat/?locale-attribute=en) has been authorized by the titular of the intellectual property rights **only for private uses** placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized neither its spreading nor availability from a site foreign to the UPCommons service. Introducing its content in a window or frame foreign to the UPCommons service is not authorized (framing). These rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

An approach for hydrogeological data management, integration and analysis

Plataforma de Gestió i Anàlisi de dades geològiques e hidrogeològiques a la ciutat de Barcelona

Rotman A. Criollo Manjarrez

Supervisors: Enric Vázquez-Suñé
Violeta Velasco

Barcelona Cicle de l'Aigua, S.A. (BCASA)
Universitat Politècnica de Catalunya (UPC)
Institute of Environmental Assessment and Water Research (IDÆA), CSIC

PhD Thesis

January 2019

The author of this thesis thanks the financial support from the Catalan Industrial Doctorates Plan of the Secretariat for Universities and Research, Ministry of Economy and Knowledge of the Generalitat de Catalunya (AGAUR); the Barcelona Water Cycle ("Barcelona Cicle de l'Aigua, S.A.", BCASA), from the Barcelona city Council; and the Spanish Council for Scientific Research (IDAEA-CSIC).

Cover design: Angels Fernández-Mora
Background cover map: The background cover map has been customised from the map of the Barcelona area and the project for its improvements and enlargement (Ildefons Cerdà i Sunyer, 1859). Source: Museu d'Historia de la Ciutat, Barcelona. Public domain.

Abstract

The conceptualisation of a groundwater system involves continuous monitoring and evaluation of a large number of parameters (*e.g.*, hydraulic parameters). Regarding hydraulic properties of the aquifers, their quantification is one of the most common problems in groundwater resources and it is recognised that all methods to obtain them have their limitations and are scale dependants. Therefore, it is necessary to have methods and tools to estimate them within a spatial context and to validate their uncertainty when they are applied in an upper scale.

All these datasets collected and generated to perform a groundwater conceptual model are often stored in different scales and formats (*e.g.*, maps, spreadsheets or databases). This continuous growing volume of data entails further improving on how it is compiled, stored and integrated for their analysis.

This thesis contributes to: (i) provide dynamic and scalable methodologies for migrating and integrating multiple data infrastructures (data warehouses, spatial data infrastructures, ICT tools); (ii) to gain higher performance of their analysis within their spatial context; (iii) to provide specific tools to analyse hydrogeological processes and to obtain hydraulic parameters that have a key role in groundwater studies; and (iv) to share open-source and user-friendly software that allows standardisation, management, analysis, interpretation and sharing of hydrogeological data with a numerical model within a unique geographical platform (GIS platform).

A dynamic and scalable methodology has been designed to harmonise and standardise multiple datasets and third-party databases from different origins, or to connect them with ICT tools. This methodology can be widely applied in any kind of data migration and integration (DMI) process, to develop Data warehouses, Spatial Data Infrastructures or to implement ICT tools on existing data infrastructures for further analyses, improving data governance.

A higher performance to obtain hydraulic parameters of the aquifer has been addressed from the development of a GIS-based tool. The interpretation of pumping tests within its spatial context can reduce the uncertainty of its analysis with an accurate knowledge of the aquifer geometry and boundaries. This software designed to collect, manage, visualise and analyse pumping tests in a GIS environment supports the hydraulic parameterization of groundwater flow and transport models.

To enhance the hydraulic parameters quantification, a compilation, revision and analysis of the hydraulic conductivity based on grain size methodologies have been performed. Afterwards, the uncertainty of applying these methods on a larger scale has been addressed and discussed by comparison of the upscaling results with pumping tests.

Finally, a sharing, open-source and user-friendly GIS-based tool is presented. This new generation of GIS-based tool aims at simplifying the characterisation of groundwater bodies for the purpose of building rigorous and data-based environmental conceptual models. It allows to standardise, manage, analyse and interpret hydrogeological and hydrochemical data. Due to its free and open-source architecture, it can be updated and extended depending on the tailored applications.

Resum

La conceptualització d'un sistema hidrogeològic implica una continua monitorització i avaluació d'una gran quantitat de paràmetres (*e.g.*, paràmetres hidràulics). Pel que fa als paràmetres hidràulics de l'aqüífer, la seva quantificació és un dels problemes més comuns als estudis hidrogeològics. És àmpliament reconegut que els mètodes per obtenir aquest tipus de paràmetres tenen les seves limitacions i són dependents de l'escala d'anàlisi. Per aquest motiu, cal disposar de mètodes i eines per estimar-los dins del seu context espacial i validar la seva incertesa quan s'apliquen en una escala superior d'anàlisi.

Les dades recopilades i generades per realitzar un model conceptual hidrogeològic sovint s'emmagatzemen en diferents escales i formats (*e.g.*, mapes, fulls de càlcul o bases de dades). Aquest volum de dades en continu creixement requereix d'eines i metodologies que millorin la seva compilació i gestió per al seu posterior anàlisi.

Les contribucions realitzades en aquesta tesi son: (i) proporcionar metodologies dinàmiques i escalables per migrar i integrar múltiples infraestructures de dades (infraestructures de dades espacials i no espacials, o la implementació d'eines TIC); (ii) obtenir un major rendiment de l'anàlisi hidrogeològic tenint en compte el seu context espacial; (iii) proporcionar eines específiques per analitzar processos hidrogeològics i obtenir paràmetres hidràulics que tenen un paper clau en els estudis d'aigües subterrànies; i (iv) difondre software de codi lliure i de fàcil accés que permeti l'estandardització, gestió, anàlisi, interpretació i intercanvi de dades hidrogeològiques amb un model numèric dins d'una única plataforma de informació geogràfica (SIG).

S'ha dissenyat una metodologia dinàmica i escalable per harmonitzar i estandarditzar múltiples conjunts de dades de diferents orígens, o bé per connectar aquestes infraestructures de dades amb eines TIC. Aquesta metodologia pot ser implementada en qualsevol tipus de procés de migració i integració de dades (DMI), per a desenvolupar infraestructures de dades espacials i no espacials, o bé per implementar eines TIC a les infraestructures de dades existents per a anàlisi addicionals; millorant així la governança de les dades.

Un major rendiment per obtenir els paràmetres hidràulics de l'aqüífer s'adreça des del desenvolupament d'una eina SIG. La interpretació dels assaigs de bombament dins del seu context espacial, pot reduir la incertesa del seu anàlisi amb un coneixement precís de la geometria i els límits de l'aqüífer. Aquest software dissenyat per recopilar, administrar, visualitzar i analitzar els assaigs de bombament en un entorn GIS, dóna suport a la parametrització hidràulica dels models de flux i transport d'aigües subterrànies.

Per millorar la quantificació dels paràmetres hidràulics, es va realitzar una compilació, revisió i anàlisi de la conductivitat hidràulica basada en metodologies de mida de gra. Posteriorment, s'ha considerat i discutit la incertesa d'aplicar aquests mètodes en una escala major comparant els resultats de la millora d'escala amb les proves de bombament.

Finalment, es presenta una eina SIG lliure, de codi obert i de fàcil aplicació. Aquesta nova generació d'eines SIG pretenen simplificar la caracterització de les masses d'aigua subterrània amb el propòsit de construir models conceptuals ambientals rigorosos. A més, aquesta eina permet estandarditzar, gestionar, analitzar i interpretar dades hidrogeològiques i hidroquímiques. Donat que la seva arquitectura és de codi lliure i obert, es pot actualitzar i ampliar segons les aplicacions personalitzades que cada usuari requereixi.

Resumen

La conceptualización de un sistema hidrogeológico implica el continuo monitoreo y evaluación de una gran cantidad de parámetros (*e.g.*, parámetros hidráulicos). Con respecto a los parámetros hidráulicos, su cuantificación es uno de los problemas más comunes en los estudios hidrogeológicos. Es ampliamente reconocido que los métodos para obtener este tipo de parámetros tienen sus limitaciones y son dependientes de la escala de análisis. En este sentido, es necesario disponer de métodos y herramientas para estimarlos dentro de su contexto espacial y validar su incertidumbre cuando se aplican en una escala superior de análisis.

Los datos recopilados y generados para realizar un modelo conceptual hidrogeológico a menudo se almacenan en diferentes escalas y formatos (*e.g.*, mapas, hojas de cálculo o bases de datos). Este volumen de datos en continuo crecimiento requiere de herramientas y metodologías que mejoren su compilación y gestión para su posterior análisis.

Las contribuciones realizadas son: (i) proporcionar metodologías dinámicas y escalables para migrar e integrar múltiples infraestructuras de datos (ya sean infraestructuras de datos espaciales y no espaciales, o la implementación de herramientas TIC); (ii) obtener un mayor rendimiento del análisis hidrogeológico teniendo en cuenta su contexto espacial; (iii) proporcionar herramientas específicas para analizar procesos hidrogeológicos y obtener parámetros hidráulicos que desempeñan un papel clave en los estudios de aguas subterráneas; y (iv) difundir software de código abierto y de fácil acceso que permita

la estandarización, gestión, análisis, interpretación e intercambio de datos hidrogeológicos con un modelo numérico dentro de una única plataforma de información geográfica (SIG).

Se ha diseñado una metodología dinámica y escalable para armonizar y estandarizar múltiples conjuntos de datos de diferentes orígenes, o bien para conectar éstas infraestructuras de datos con herramientas TIC. Esta metodología puede ser implementada en cualquier tipo de proceso de migración e integración de datos (DMI), para desarrollar infraestructuras de datos espaciales y no espaciales, o para implementar herramientas TIC en las infraestructuras de datos existentes para análisis adicionales; mejorando así la gobernanza de los datos.

Un mayor rendimiento para obtener los parámetros hidráulicos del acuífero se ha abordado desde el desarrollo de una herramienta SIG. La interpretación de ensayos de bombeo dentro de su contexto espacial, puede reducir la incertidumbre de su análisis con un conocimiento preciso de la geometría y los límites del acuífero. Este software diseñado para recopilar, administrar, visualizar y analizar las pruebas de bombeo en un entorno SIG, apoya la parametrización hidráulica de los modelos de flujo y transporte de aguas subterráneas.

Para mejorar la cuantificación de los parámetros hidráulicos, se ha realizado una compilación, revisión y análisis de la conductividad hidráulica basada en metodologías de tamaño de grano. Posteriormente, se ha considerado y discutido la incertidumbre de aplicar estos métodos en una escala mayor comparando los resultados de la mejora de escala con los obtenidos en ensayos de bombeo.

Finalmente, se presenta una herramienta SIG libre, de código abierto y de fácil aplicación. Esta nueva generación de herramienta SIG pretende simplificar la caracterización de los cuerpos de agua subterránea con el propósito de construir modelos conceptuales ambientales rigurosos. Además, esta herramienta permite estandarizar, gestionar, analizar e interpretar datos hidrogeológicos e hidroquímicos. Gracias a su arquitectura de código libre y abierto, se puede actualizar y ampliar según las aplicaciones personalizadas que cada usuario requiera.

Acknowledgements

I appreciate and am very grateful to all who gave me their support along these years.

Doy las gracias a todas/os las/os que me han apoyado durante estos años.

En primer lugar doy las gracias a Enric y Violeta por la confianza depositada en mí y por la ayuda recibida.

A mis compañeras/os y amigas/os de BCASA: Xavi, Silvia, Mª José, Angles, Monica, Fidel, Juan Antonio, Maria, Ariadna, Dani, Lina, Raquel, Rosa, Ramón, Montse, Alberto, Jose y demás compañeros.

A Carlos por sus consejos, por sus comentarios y ayuda durante esta tesis.

A mi familia, que sin su historia y todo lo que hemos pasado, yo no habría escrito esta tesis. Esta tesis está dedicada a ellas y ellos, pero sobretodo a mis abuelos: Inés, Elvia, Filito, Tránsito, Segundo Isaías y Wilson.

A todo el GHS, por el cariño y por todo lo compartido.

A mis amigas/os, que también han caminado junto a mí durante estos años: Miriam, Arafat, Alex, David, Quim, Laura, Nuria, Maricarmen, Raúl, M.A., Jordi, Albert, Silvia, Gerard, Paco, Jean, Víctor, Iñijo, Marta, Mireia, Bayra, Juancho, Geert, Marcela, Isaac, Loreto, Leo, Hamideh, Felipe, Mar, Gonzalo, Juan, Estanis, Tybaud, Alba, Sonia, Nabil, Jing Jing, Berta, Anna, Yoar, Cris, Carme, Alessandro, Tomais, Alec, Dennis, Alex, Lazaro, Lidia, Robert .... seguro que me dejé tu nombre. Nos vemos en el café y me dices (que hoy hay dulce para acompañarlo). Echo mucho de menos las discusiones y campañas

de campo con Alex y Laura. Gracias por el apoyo y empuje durante todos estos años!

Y mil gracias Angels, gracias por tus consejos, ayuda, apoyo, confianza, los paseos divagando del mundo e intentándolo arreglar (de momento sin éxito... creo que tendremos que hacer más paseos juntos). Pero sobretodo mil gracias por el amor que me has dado.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation and objective

Groundwater is the greatest resource of fresh water in the world (Rossetto et al., 2018). Its spatio-temporal distribution and quality are fundamental variables of the hydrological cycle (Gleeson et al., 2016a). Moreover, groundwater bodies have different anthropic uses from water and energy supply to CO2 storage. Hence, controlling the behaviour of the groundwater system in order to ensure their sustainability has become of paramount importance in the water and energy management. The conceptualization of a groundwater system involves continuous monitoring and evaluation of a large number of parameters such as groundwater levels, temperature, pH, or hydraulic parameters. Regarding hydraulic properties, their quantification is one of the most common problems in groundwater resources (Bear, 1979; Uma et al., 1989). They are essential to understand the spatial variation of the flow regime restricting, for instance, the movement of pollutants through the aquifer. Hydraulic parameters are conditioned by the aquifer geometry and their hydraulic boundaries (*e.g.*, rivers, faults, dewaterings, tunnels). It is recognised that all methods to obtain them have their limitations (Chapuis et al., 2012; Ren et al., 2016) and are scale dependents (Sánchez-Vila et al., 1996;

Vukovic and Soro, 1992). Nevertheless, their application in scales beyond their limitations is usually required because of the information available (Gleeson et al., 2016a). Therefore, it is necessary to have methods and tools to estimate them within a spatial context and to validate their uncertainty when they are applied in an upper scale.

All these datasets collected and generated to perform a groundwater conceptual model can be reinforced by additional information (such as geology or isotopes). This information is often stored in different scales and formats (*e.g.*, maps, spreadsheets or databases). This continuous growing volume of data entails further improving on how it is stored. This is one of the main issues lead by Information Technology (IT) (Yue and Tan, 2018). Specifically, Information and Communication Technology (ICT) has had a significant impact on how these data are compiled, shared and, consequently, how these data are improving the decision making process in the last years (Foglia et al., 2017). ICT tools not only facilitate the dissemination of information, they enable a more inclusive and participatory approach (https://www.ict4water.eu). Thus, ICT tools may provide a framework to avoid conflicts of interests as they include stakeholders in the participation of decision processes (Korres and Schneider, 2018).

Location of data, knowing also as geospatial data, spatial data or geodata (Goodchild, 2018), use to be conceptualised as a collection of geometric and alphanumerical data. Currently, Geographic Information Systems (GIS) are suitable tools developed in order to facilitate their visualisation and their manipulation (*e.g.*, Goodchild, 2010, 2018; Steyaert and Goodchild, 1994; Sjoukema et al., 2017). Furthermore, GIS platforms aids to understand spatial patterns in data and identify process.

The use of GIS platforms together with ICT tools (*i.e.*, GIS-based tools) aid stakeholders to solve problems within their spatial context (Bojórquez-Tapia et al., 2001; Korres and Schneider, 2018; Foglia et al., 2017; De Filippis et al., 2018). Additionally, GIS-based tools may facilitate to perform a systematic methodology of analysing geodata. This systematic methodology, thus, facilitates data management in a more efficient way and, at the same time, can validate the quality of data and allow their reuse. To guarantee the reuse of all data related to groundwater, and to store them in a harmonised way are required spatial databases. Spatial databases are builded from a conceptual data structure where geodata have necessarily to simplify the reality.

Cabalska et al. (2005), Gogu et al. (2001), Strassberg (2005), Velasco (2013), Wojda et al. (2013) are some examples of spatial data structures related to groundwater that ensure a proper sharing of knowledge. They consist of GIS-based tools that focus on storing, managing, visualising and analysing hydrogeological data. In addition, they may provide information to create input files to numerical models using the common tools available in GIS platforms.

Despite these significant advancements, there is still the need of new developments: (i) to adapt these tools to specific institutions and / or third-party databases; (ii) to ensuring data validation; (iii) to analysing further hydrogeological processes and obtaining aquifer parameters from field measurements; and (iv) to enabling the groundwater community to use these platforms.

In the light of last discussion and requirements above mentioned, the aims of this thesis are designing, developing and applying methods and tools to improve management, migration, integration, analysis and interpretation of hydrogeological data. To achieve these objectives, specific goals of this dissertation are:

i to provide dynamic and scalable methodologies for migrating and integrating multiple infrastructures (data warehouses, spatial data infrastructures, ICT tools);

ii to gain higher performance of their analysis within their spatial context;

iii to provide specific tools to analyse hydrogeological processes and to obtain hydraulic parameters that have a key role in groundwater studies; and iv to share open-source and user-friendly software that allows standardisation, management, analysis, interpretation and sharing of hydrogeological data with a numerical model within a unique geographical platform.

1.2 Thesis Outline

The thesis has been divided in six chapters. After this introductory chapter, next chapters have been written addressing the specific goals of this thesis.

Chapter 2 describes a dynamic and scalable methodology to harmonise and standardise multiple datasets and third-party databases from different origins in a single spatial infrastructure. Its application in the Barcelona City was performed by developing frameworks of migration and integration among the Barcelona City Council systems and other external information obtained from scientific developments and projects related with hydrogeological and civil works studies.

In next chapters, 3 and 4, are addressed and described specific tools to analyse hydrogeological processes and to obtain hydraulic parameters in several scenarios.

Chapter 3 describes the design, structure, development and application of GIS-based tools for collecting, managing, analysing, processing and interpreting data derived from pumping tests. The interpretation within its spatial context of a hydraulic test can reduce the uncertainty with an accurate knowledge of the aquifer geometry and boundaries. Their interpretation can be stored in the same database in order to increase the knowledge of the study area. Chapter 4 reviews and analysis compiled equations to obtain the hydraulic conductivity from grain size data. The uncertainty of applying these methods on a larger scale has been addressed and discussed by two methods of upscaling and their comparison with pumping tests performed in the study area.

Chapter 5 gives an overview of an open-source and user-friendly GIS-based tools that allow standardisation, management, analysis and interpretation of hydrogeological and hydrochemical data. This new generation of tools can build input files (*i.e.*, surfaces of hydrogeological units) for numerical models within a unique geographical platform. Due to its open-source architecture, it can be updated and extended depending on the tailored applications.

Finally, Chapter 6 synthesises the whole research. It consists of a summary of the issues addressed in this thesis with a discussion of what are the new contributions to the scientific and technical community, and what are the future improvements.

1.3 Framework of this thesis

This thesis has been developed in the framework of the Industrial Doctorates Plan ("Doctorat Industrial", DI). *"The aim of the Industrial Doctorates Plan is to contribute to the competitiveness and internationalisation of Catalan industry, strengthen the tools for recruiting the talent generated in the country and place future PhD holders in the right place to carry out R&D&I projects"* (AGAUR, 2019).

The collaboration agreement between the Barcelona Water Cycle ("Barcelona Cicle de l'Aigua, S.A.", BCASA), from the Barcelona city Council; the Universitat Politècnica de Catalunya (UPC) and the Spanish Council for Scientific Research (IDAEA-CSIC) is summarised below.

Summary of the DI project

The increase in the population is conducted by the increase in exploitation of natural resources. Particularly in developing areas where population and industrial growth is much more known and, therefore, its demand of water as well. This population growth requires large infrastructure to meet their needs and many of them interact with the subsoil. It implies an alteration of the territory that increase the pollution vulnerability of the subsoil (gas stations, waste from mines, ports, hospitals, landfills, etc.) that could be prevented. Currently, new techniques and methods have been developed to improve their effects and thus guarantee a good state of the environment, still undergoing in an improvement stage.

Water resources management presents an added challenge: climate change. This produces variations of precipitation and temperatures mainly; which can lead to a decrease or increase in the levels of groundwater. Consequently, these climatic variations cause that the demand of present and future water (and its quality) are more complex to quantify. The evaluation and quantification of the magnitude of these climate changes are necessary to perform a proper management of these resources.

Better management and reduction of these problems is only possible knowing in more detail the processes involved in the renovation of waters (hydrogeochemical processes). Thus it is possible to guarantee a proper quantity and quality of water for current and future demand. For this reason, hydrogeological models are essential. These models allow conceptualising and quantifying the hydrogeological processes simulating various scenarios such as droughts, exploitation of resources, evolution of quality, behaviour of pollutants in the hydrogeological system, interaction of underground works conditioned by the hydraulic and geomechanical behaviour of the terrain, etc.

The modelling of hydrogeological processes within the heterogeneous environment is very complex and often they give high degrees of uncertainty in their results. The intrinsic natural heterogeneity of the sedimentary environment and the poor development of the tools to manage the available data, regarding to the visualisation of data, are not enough suitable for end users. These types of models often show the 2D environment, although the natural heterogeneity in the natural environment represent a complex three-dimensional distribution; so it is essential to upgrade and display all the data that are available in the same 3D environment.

These problems have been studied for years, giving basically solutions to the most immediate problems, generating a large volume of data that is continuously increasing. That is why good management of data is very important to: (i) obtain some quality and reliable data; (ii) perform the analysis more quickly; and (iii) optimise decision making.

The main objective is to propose solutions to these problems so that they can be carried out easily by researchers and professionals to be applied for the management of the groundwater in the city of Barcelona. To achieve this aim, the following points are proposed: (i) improve the management and exploitation of data tools, developing tools and methodologies that facilitate the integration of 3D hydrogeological models; (ii) to be able to represent more accurately the heterogeneous environment that surrounds us to apply them within the 3D hydrogeological models and (iii) apply them to real problems existing in the Barcelona area to evaluate and quantify the improvements.

All of this results in greater protection of the city's water resources, both in forecast (management rules in the face of possible scenarios, including climate change scenarios) and pollution (study and prediction of possible pollution or marine intrusion), thus affecting a good quality and quantity of water for the current demand and being able to safely increase future demand; reducing the economic and environmental costs of underground works.

Object of the Agreement

This Agreement aims to establish the collaboration regime between the parties for the development of the research project and the industrial doctorate "Plataforma SIG de gestió y análisis de dades geològiques e hidrogeològiques a la ciutat Barcelona", the content which will become the object of this doctoral thesis.

People in charge of the project management

The direction of the research project will be carried out by people from the following academic and business environment.

Academic environment

Institute of Environmental Assessment and Water Research (IDAEA), the Spanish Council for Scientific Research (CSIC) together with The Universitat Politècnica de Catalunya (UPC).

Academic Project Manager:

Enric Vázquez Suñé. Vice director of the IDAEA-CSIC.

Supervisors:

Enric Vázquez Suñé, PhD. (IDAEA-CSIC)

Violeta Velasco Mansilla. PhD. Specialist in Spatial databases and GIS

Institute of Environmental Assessment and Water Research (IDAEA), the Spanish Council for Scientific Research (CSIC).
C/ Jordi Girona, 18, 08034, Barcelona.

Advisor:

Xavier Sánchez Vila, Professor

The Department of Civil and Environmental Engineering (DCEE), Mòdul D2, Campus Nord, UPC. C/ Jordi Girona,1 -3, 08034 Barcelona.

Business environment

Barcelona Water Cycle ("Barcelona Cicle de l'Aigua, S.A.", BCASA), Barcelona city Council. C/ Acer, 16 08038 Barcelona.

Business Project Manager:

Xavier Varela i Alegre. Head of Planning and Innovation

Contributors:

Sílvia Burdons Cercós. Head of Cartography Service

Mª José Chesa Marro. Head of Foreign and R&D Services

1.4 Scientific and technical production

Some scientific and technical deliverables have been done in the framework of this thesis. They are the result of interaction among scientific and technical entities and of synergies among them. This interaction, at the same time, helped to bringing insights on the needs of public and private entities. Their result is shown in Appendix A of this document. This appendix is divided by scientific publications directly related with this thesis and collaborations in other scientific contributions, chapters in books, proceedings in congresses, registers/patents of software, and technical reports in which the author has been committed and involved actively during this thesis.

Chapter 2

On hydrogeological data migration and integration

Summary

Data infrastructures (for example, data warehouses and spatial data infrastructures) and their governance are continuously being used and developed because of their impact on groundwater management. In addition, these kinds of data infrastructures currently require information communication technology (ICT) tools for improving geodata governance by delivering quality information to users and helping them perform further analyses on a unique platform. A key factor in this trend is the need to reduce complexity in data migration and integration (DMI) projects and processes. To implement these systems (develop a data infrastructure or connect it to ICT tools), facilitation and optimisation of DMI is required.

Based on: Criollo, R., Vázquez-Suñé, E., Cardona, F., Burdons, S., Enrich, M. On hydrogeological data migration and integration. Submitted in Environmental Modelling & Software (2019).

We present a novel methodology to facilitate and optimise DMI to enhance data governance. The implementation of this novel DMI methodology facilitates merging multiple sources of information, installing new systems to exploit stored information while using the original systems of information storage, and upgrading databases, formats or standards that may not be supported in the future to one that is supported or most appropriate, among other processes. The outcomes of its application by the Barcelona City Council (Spain) are used to optimise the analysis process and share their results on the hydrogeological status across the city. DMI models performed for Barcelona systems can be easily adapted to other external datasets, increasing the volume of quality data to improve the understanding of the groundwater system behaviour and the monitoring network in the city. Other benefits observed during implementation of this DMI methodology include improved data quality checking in the original sources of information and encouraging collaboration among departments related to groundwater bodies. The proposed methodology can be widely implemented in any kind of DMI project to develop data infrastructures or to implement ICT tools for further analyses, improving groundwater management.

2.1 Introduction

Groundwater resources have various anthropic uses, from water and energy supply to CO_2 storage. The impacts of these uses on the behaviour and quality of groundwater must be determined to guarantee the presence of long-term sustainable resources. Groundwater studies consist of constructing a conceptual model of the system supported with data available and then performing experiments to verify the understanding of the system behaviour (*i.e.*, across field, laboratory, and numerical model results). Legacy or historical data play a key role in understanding the baseline of the subsurface system and its evolution. Multiple sources of information are commonly compiled during a groundwater study that is not always stored within a consistent nomenclature since the information usually has not been collected for a common purpose

(Sundell et al., 2016); despite the increasing groundwater standards (*e.g.*, INSPIRE, 2011; OGC, 2011; and GWML, Boisvert and Brodaric, 2012). The time and resources spent collecting, understanding and analysing (*e.g.*, quality data) datasets are important for gaining confidence in the results obtained. That effort, typically, is not taken into account when a project plan is in development. For hydrogeological studies using historical and current datasets, the interpretations and results are commonly ignored after the study is completed. Therefore, after the project, in general, this information is not available to use in other projects and is not accessible to other people, increasing the costs of projects, resulting in decisions based on less information and thus reducing groundwater governance.

For better groundwater governance, it is convenient to implement effective data governance (*e.g.*, ISO/IEC 38500:2008). One of the pillars of effective data governance is data management (Brown, 1997; Otto, 2011; Wende, 2007). Data management is successful when data are harmonised, collected, structured, stored, available, understandable, reusable, and checked for errors, ensuring that the data model can be maintained (Carleton et al., 2005; Fitch et al., 2016; Horsburgh et al., 2009; Kao et al., 2011; Ranatunga et al., 2011; Tartar, 2008). These aspects ensure data integrity, avoid loss or duplication of raw data (Altheide, 2008; Wortman, 1992) and reduce uncertainty in the spatial and non-spatial data quality (Guptill and Morrison, 1995).

Currently, there are several developments in stewardship hydrogeological information that can be implemented in industry, government and academia to reduce the time and resources needed to perform hydrogeological analyses. These kinds of data infrastructures, data warehouses and/or spatial data infrastructures (SDIs) (*e.g.*, Andreadis et al., 2017; Blodgett et al., 2011; Kingdon et al., 2016; Luo et al., 2011; Pasquale et al., 2008; Mazzei and Di Guida, 2018) and their governance are developing worldwide because of their importance in ensuring resource sustainability (Sjoukema et al., 2017). Data infrastructures and governance have many definitions (*e.g.*, Hendriks et al., 2012, Phillips et al., 1999, Korhonen et al., 2013, and Otto, 2011, and Wende, 2007, respectively) and could be very different in the future (Sjoukema et al., 2017).

The implementation of information communication technology (ICT) tools in data warehouses or SDI systems is being required more often. In fact, the OGC highlights the importance of integrating different tools to give integrative solutions of different processes. These connections represent a big step forward for geodata governance through delivering quality information to users and helping them in further analyses. Hence, data infrastructures and ICT tools reinforce the decision process for an optimal groundwater stewardship in a unique platform.

Rossetto et al., 2018; Strassberg, 2011; Velasco, 2013; and Yang and Lin, 2010 are examples of current GIS-based ICT tools related to groundwater management. These tools have specific structures to store, visualise and create other spatial analysis and numerical models. To implement any GIS-based ICT tools in any data systems, there are three possible options:

- The first option is related to adapt the external tools to the user‘s data structure. This option could be difficult to execute and produce uncertain results due to the alteration of internal code; typically, we may have not detailed information about the code and its internal structure. In addition, most of the end user’s (mainly hydrogeologist) do not have properly skills to achieve it.

- The second option is to adapt the user‘s data structures to the data required by specific tools; this may be a more feasible solution, but it requires high costs in terms of time and resources. This option is often taken by teams that do not work in the same field, and the data systems may not have well documented. Additionally, the lack of open source codes makes more difficult in achieve this task and there is a risk of losing information and/or communication among our systems with other external tools, such as connection to monitoring sensors.

- A third option is to migrate and integrate information from internal systems to connect them with the ICT tools required without modifying the internal structures of both the source and destination data systems. This solution enables us to:

i avoid wasting time due to changing the old systems,

ii improve the coordination among work teams,

iii boost reluctant users to use new technologies, and (iv) provide minimal impact on ICT, reducing cost and time.

Other issues to be considered related with data migration and integration (DMI) are the following: (i) to upgrade databases or formats that may not be supported in a future to one that is supported or most appropriate (*e.g.*, sensors, rasters, databases); (ii) to manage spatial or raster (ESRI, 2018); (iii) to merge external sources of information from the original user's data systems; and (iv) to install new systems to exploit data stored (data marts), while continue using original systems to store the information.

Many migration and integration tasks can be performed automatically, avoiding human errors and the waste of resources and time. However, DMI are the most critical and expensive step when new systems (data infrastructures or ICT tools) are implemented (Scheier, 2018). The quality of data migration can significantly affect business operations and intelligence, as well as the overall transaction timeline (Joy et al., 2018).

Data migration can be divided into two main types of migrations: massive (large bang migrations) and intervaled (trickle migrations) (Oracle, 2011). Data integration implies de inclusion of cleansing, reformatting and storage of data. There are many software programs, such as Safe FME (Safe Software, 2018), Kettle (Hitachi Vantara Community, 2018), Migratool (Leite et al., 2005), Talend (Talend, 2006), Oracle (Oracle, 2011), among others, that perform DMI processes. Additionally, there are other specific tools to migrate from/to single formats (for more information, please see PostgreSQL), but to the best of our knowledge, in the literature, few detailed methodologies or protocols to migrate and integrate spatial information have been described (Howard, 2008; Martens et al., 2018; Rüping, 2013; Serra et al., 2018; Wagner

& Wellhousen, 2011; Wu et al., 1997), despite their demand (Goodchild, 1997; Leite et al., 2005; Tartar, 2008).

An optimal DMI is essential for decision-making with clean high-quality data, which produces a more confident and stringent groundwater governance. To ensure an optimal DMI, it is valuable to provide frameworks with which to assist and facilitate processes to connect and transform multiple data systems from different sources and formats to the required destination formats of new data systems. The connections among system structures and the organisation of a DMI model in terms of its application and maintenance should be as intuitive as possible.

Thus, there is a need to provide dynamic and scalable methodologies to migrate and integrate multiple infrastructures to improve the performance of the collected data, ensuring data quality, security and consistency as automatically as possible. Hence, at the same time, IT and groundwater skills can collaborate to improve data governance, thus improving groundwater management.

Regarding these challenges, the aim of this manuscript is to present a novel methodology to facilitate and optimise DMI processes from multiple sources and formats to new systems for further analyses to optimise groundwater management. This methodology can be widely applied for the migration and integration of information to any kind of data management and analysis platform.

After a description of the proposed methodology is given in the following section, we present an implementation of this methodology in Barcelona (Spain) and the obtained results to demonstrate its application. Finally, the methodology and the application are discussed.

2.2 Methodology

The proposed methodology is intended to migrate, transform and integrate multiple sources with different formats into a new system to improve groundwater management by achieving higher profits and increasing production. The use of the same platform for all the data related to groundwater issues encourages data sharing and cooperation among people with different skills and from different departments.

Figure 2.1 describes the main stages in a DMI process: collection of the data; definition of the strategy; execution of the model; and, finally, validation and maintenance of the DMI model and procedures.

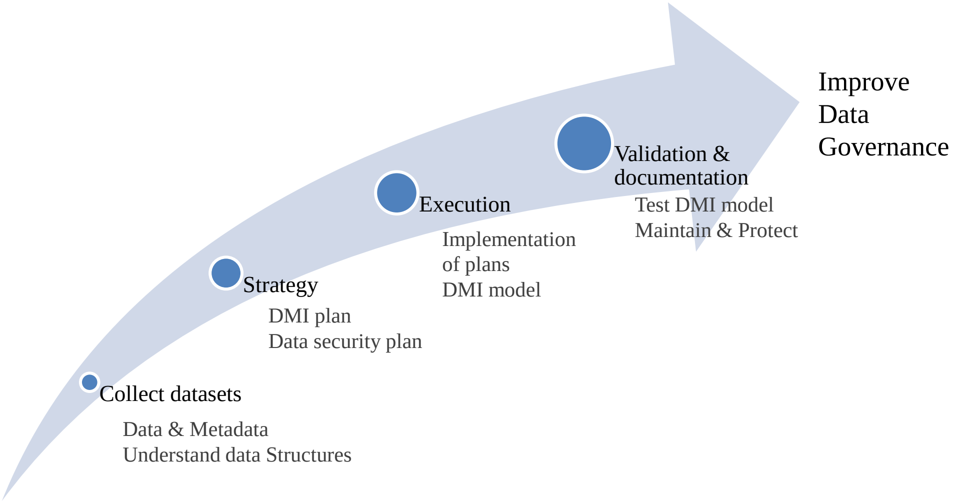


Fig. 2.1 Workflow scheme of stages to perform a DMI process.

<u>Collect datasets</u>

It is essential to meet the needs of all relevant stakeholders, from different departments or with different skills related to groundwater, to ensure proper data collection and avoid new sources of information (*e.g.*, databases from other servers, legacy spreadsheets, and raw data from sensors) during the migration and integration processes. Afterwards, it is recommended to select only the relevant information required for the new system and to solve data issues such as orphaned storage, corrupt data and/or inefficiencies. for instance, measurements without spatial positions, measurements without measurement dates, dates of measurements without measures, and redundant information.

Migrating less information ensures less error during their migration. Hence, the process of understanding data structures and their taxonomies is improved and can be used to determine a strategy for migration and integration. We refer to taxonomy as a conceptual model of how reality is simplified in a data model and is crucial to creating the structure of the data infrastructure. Managing different taxonomies increases the cost of resources, resulting in a lack of exchanged knowledge.

Strategy

Building a strategy will guide us to clarify the deadlines and objectives, reorganise the budget and check in with people playing key roles during the migration process. In addition, the strategy stage helps to create the DMI plan, create the data security plan and identify other potential issues that may occur later on.

The process of creating a proper DMI plan and its model can also be useful to avoid the overlap and/or delay of tasks during the implementation of the new system. A DMI plan consists mainly of the following steps: (i) to decide teams to lead and perform DMI processes and (ii) to decide how and when data are extracted, transformed and loaded into the new system. For instance, it is necessary to evaluate the time to perform the migration and integration processes or evaluate the type of migration (big bang migrations or trickle migrations). Trickle migration is optimal for time-dependent data, such as hydrogeological measurements, because time intervals may help users control what information is or is not migrated without interrupting the use of the original systems.

To build a strategy, a template of relationships among each source field and destination field is useful because it will guide us to a better understanding of what kind of transformations can be done.

Furthermore, in this stage, inconsistencies can be detected, such as text (string) fields that are longer in the source systems than the maximum longitude in the destination field, different time formats, different coordinate systems, different geological nomenclatures, and different hydrogeological and/or hydrochemical parameters (see Figure 2.2).

Regarding dictionaries, destination systems sometimes contain their own dictionaries that are used by applications, and the modification of their registers could be critical (*e.g.*, modify codes, nature of primary keys) for the correct functionality of external analysis software. Considering that source systems can be continued by using trickle migrations and avoiding resistance to system change, a suitable option is to create translator tables among them. That is, tables are created of the relationships among the source and destination nomenclatures to upgrade and standardise this kind of information.

Execution

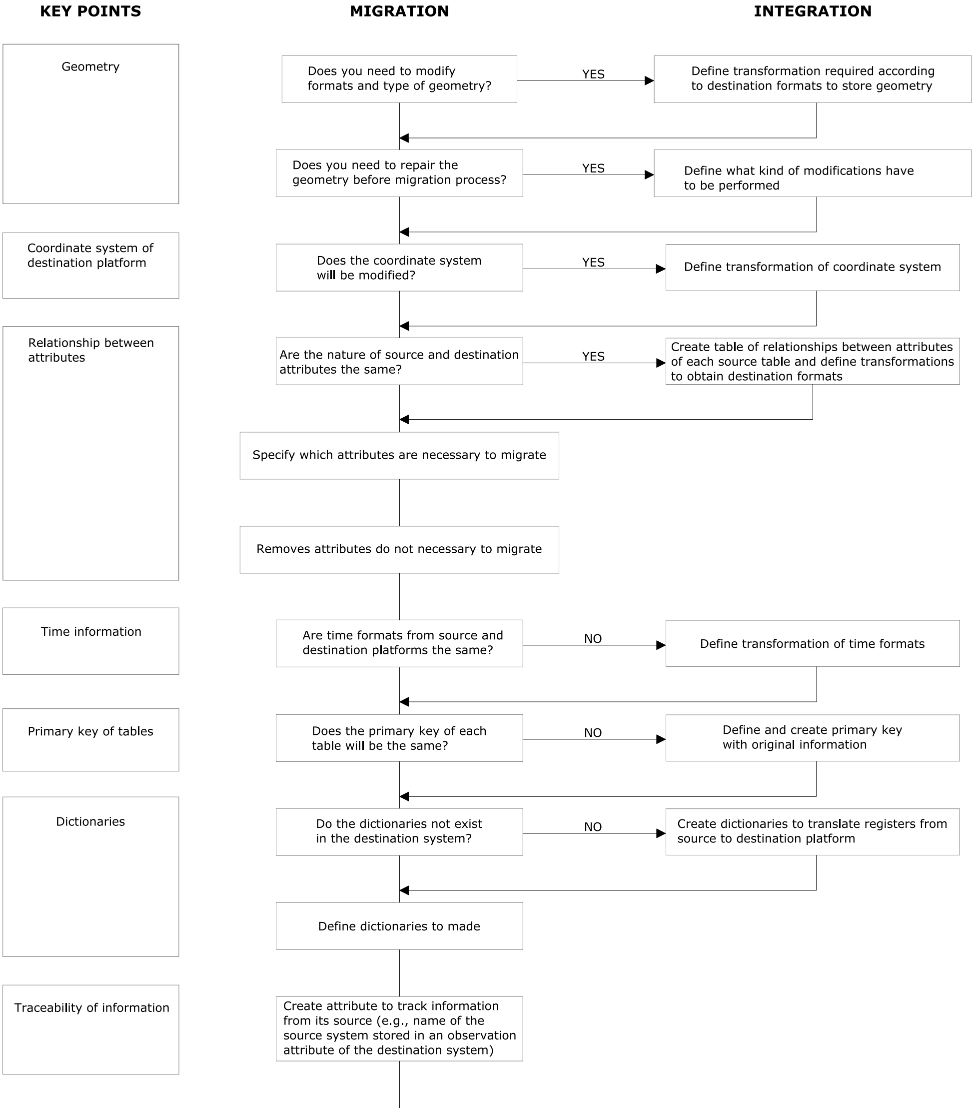
The assessment and analysis of the strategy leads to the implementation of the DMI plan by performing the DMI model. Extraction, transformation, cleansing and deduplication of data before moving it to the new destination system are the steps performed in this stage. Notably, transformations can be used as filters to check internal errors in unstructured datasets that were not checked during the data collection stage.

It is recommended to optimise the DMI model performed in order to reduce the computational time and resources. For instance, it is advised to clean up data not relevant in the destination system before performing transformations.

Validation and documentation

Validation of the strategy and the processes of extraction, transformation and information loading are essential to ensure that it is an accurate DMI model that reduces uncertainty and risk. A test of DMI processes with different sizes of geodata will create confidence in the DMI model. Execution of the migration and integration process with data control (*e.g.*, a known number of points with a known number of head registers) can be a first test.

Finally, the entire DMI model is audited and documented to maintain and protect its integrity and quality. If the DMI will utilise trickle migration without using source systems, it is recommended to follow migration and integration policies. These policies are guidelines for the DMI plan and its maintenance, which will allow for improved data governance.



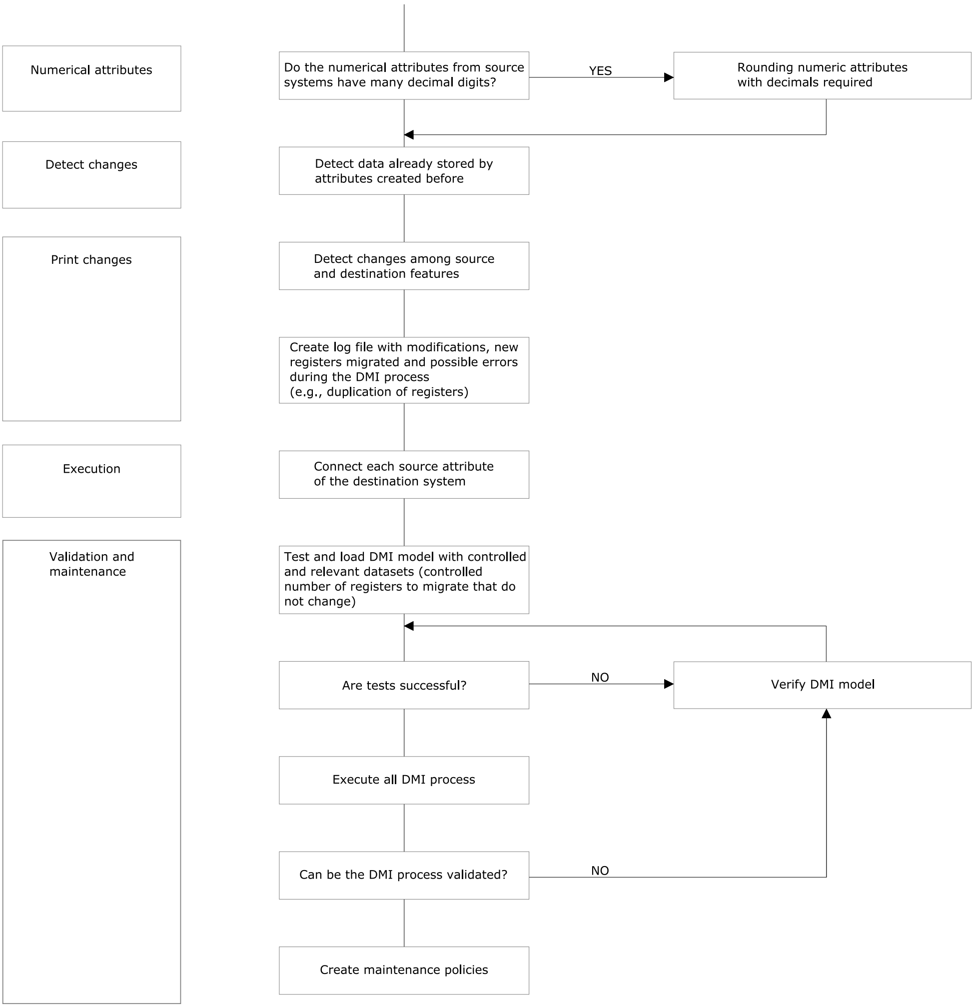


Fig. 2.2 Workflow scheme of key points and DMI steps.

2.3 Implementation by the Barcelona City Council

The application of the proposed methodology has been carried out as part of the project "GIS Platform for management and analysis of geological and hydrogeological data in the City of Barcelona" that aims to implement GIS-based ICT tools (HYDORGIS tools, Velasco, 2013) to improve groundwater management in the city, facilitating data standardisation, visualisation, analysis and interpretation. Additionally, this GIS-based ICT tool aims to facilitate and improve the reporting of results of groundwater quality and quantity by common hydrochemical and hydrogeological diagrams by creating input files to perform numerical models for further analysis (see Figure 2.3).

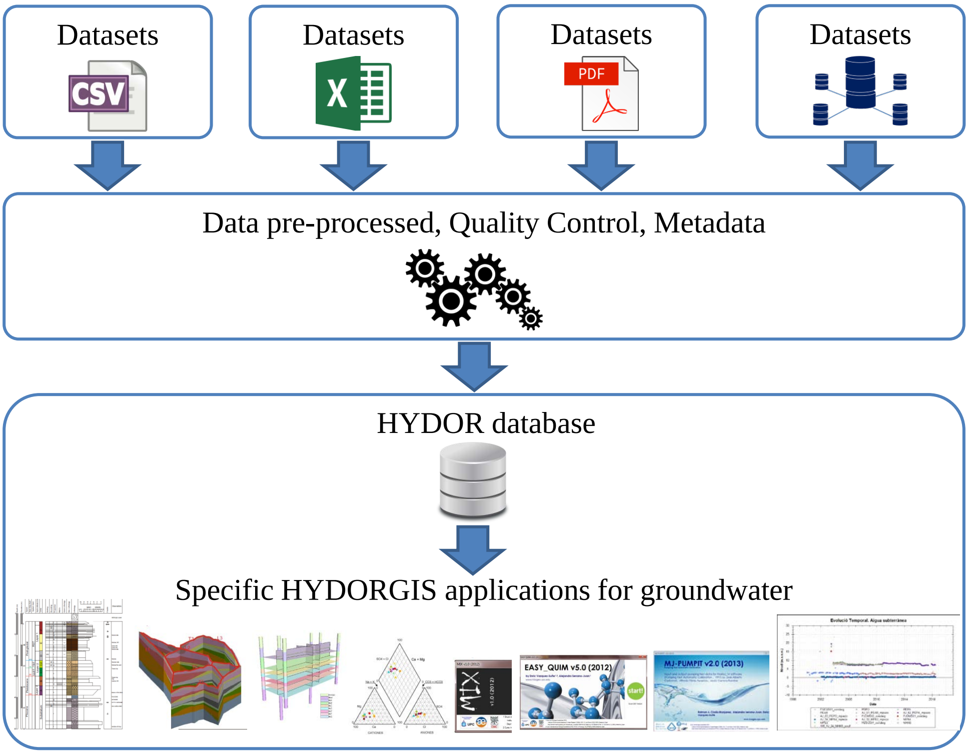


Fig. 2.3 Workflow scheme of migration and integration of multiple sources into HYDORGIS tools.

Hydrogeological information stored in the Barcelona City Council (Barcelona water cycle, BCASA) systems and other information collected from other sources, such as scientific investigations and civil works performed over multiple years in the city, were migrated and integrated into GIS-based tools.

The core of HYDORGIS tools is the HYDOR spatial database (Velasco, 2013), where all the information related to hydrogeology is stored. Hence, HYDORGIS tools query its information to display common hydrogeological diagrams and maps. There are five toolbars for controlling the main hydrogeological information related to: geology (HEROS, Velasco et al., 2012a; HEROS3D, Alcaraz, 2016b), hydrochemistry (QUIMET, Velasco et al., 2014), hydrogeology (HYYH, Criollo et al., 2016) and environmental modelling (ArcArAz, Alcaraz, 2016). The workflow for performing a hydrogeological analysis using HYDORGIS tools is graphically described in Figure 2.4.

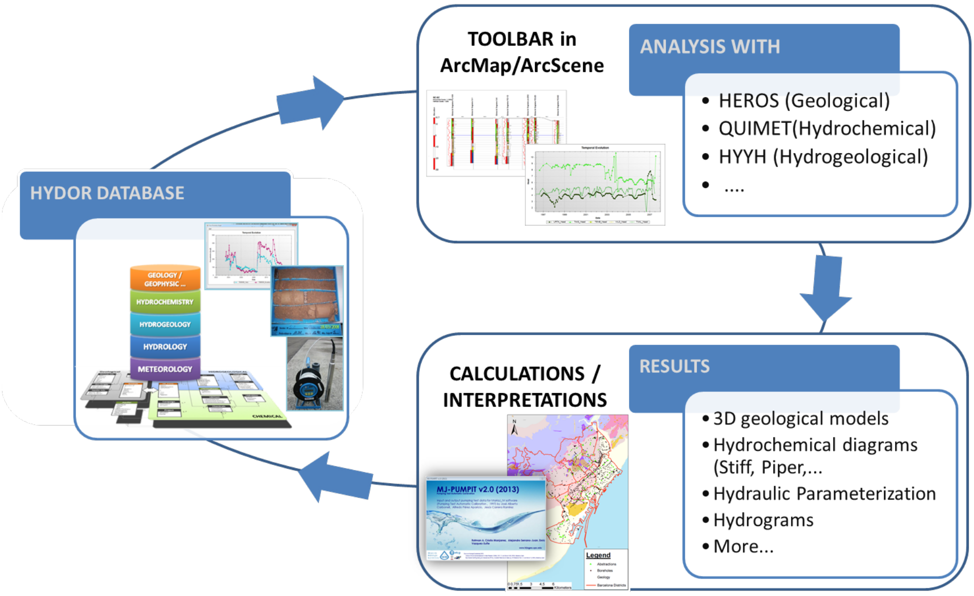


Fig. 2.4 Scheme of the workflow using HYDORGIS tools to analyse data stored in the HYDOR database, and how their interpretations are stored in the database.

Implementation

Collect datasets

Apart from the BCASA systems, several other sources of information were used (see Figure 2.5), namely, scientific field campaigns and external information from civil works performed in the city.

Scientific field campaigns and information from civil works

Information collected from several scientific field campaigns and civil works performed in Barcelona have different formats and taxonomies.

More than 4800 points (*i.e.*, from boreholes, wells, piezometers, rivers and springs, were collected and homogenised in the HYDORGIS tools (Figure 2.5). Their measurements and interpretations have been collected from different origins with different formats, sometimes unstructured formats.

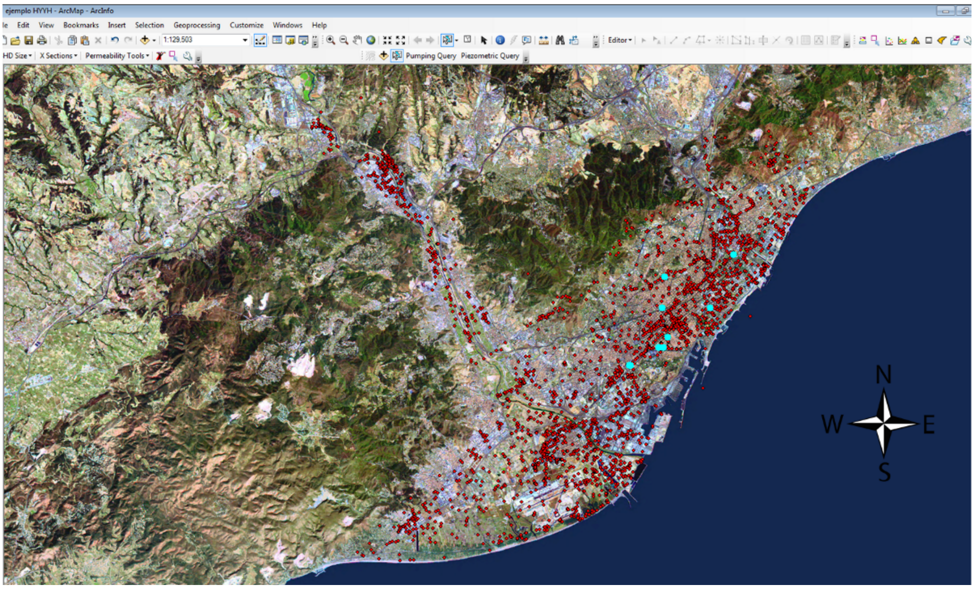


Fig. 2.5 Illustration of points where hydrogeological information has been collected and stored in the HYDOR database.

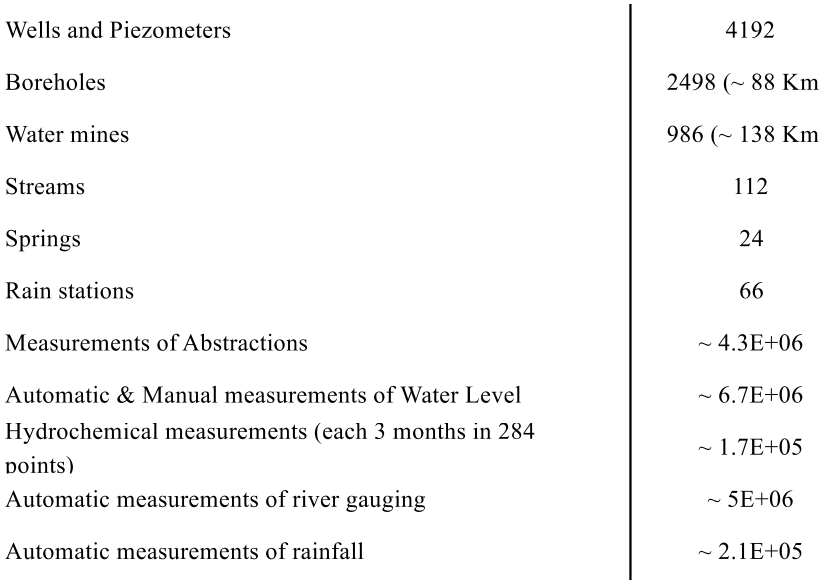
All of these data were applied in scientific studies; for instance, to create a 3D geological model of Barcelona (Vázquez-Suñé, 2016a).

Barcelona water cycle systems

BCASA has three main databases related to groundwater structured in Smallworld (called SITE) and in Oracle (SIEC and SICA). SITE stores geospatial information, while SIEC and SICA store hydrogeological and hydrochemical alphanumerical information, respectively.

Smallworld (GE Solutions, 2018) is a GIS object-oriented, database-driven product that provides architecture at the heart of many applications, such as those used for planning electric, gas and water distribution systems, designing telecommunication networks and evaluating strategic market opportunities (GE Solutions, 2018).

The relation among alphanumerical geospatial information is made by indexes. Table 2.1 shows the number of hydrogeological datasets stored in BCASA systems, which are continuously growing.



Wells and Piezometers	4192
Boreholes	2498 (~ 88 Km)
Water mines	986 (~ 138 Km)
Streams	112
Springs	24
Rain stations	66
Measurements of Abstractions	~ 4.3E+06
Automatic & Manual measurements of Water Level	~ 6.7E+06
Hydrochemical measurements (each 3 months in 284 points)	~ 1.7E+05
Automatic measurements of river gauging	~ 5E+06
Automatic measurements of rainfall	~ 2.1E+05

Table 2.1 Typology of points stored in the BCASA systems related to groundwater.

Strategy

After understanding and collecting all the necessary information, the strategy to perform DMI processes was divided into 2 parts: (i) massive migration of all the relevant information for the groundwater analysis until one day before the date on which this transfer is made and (ii) trickle migration of new data as of the last date of migration. The last type of migration by intervals, trickle migration, will be controlled by the time interval desired by the user and needs a verification of the status of the records already stored in HYDOR. Therefore, records included in HYDOR (imported during the massive or previous trickle migration) will be compared with registers that exist in the BCASA systems. In addition, this comparison will be useful to control possible anomalies of the piezometric levels recorded by the operating department. These anomalies are irregular records that occur for different reasons, *e.g.*, variations in the height of the sensors and malfunction of the device. If anomalies are detected, piezometric levels would be modified; then, records in HYDOR also have to be modified due to this comparison of information stored in both the original and destination sources.

To homogenise all of this information, the methodology proposed in this manuscript has been applied to introduce new data into the HYDORGIS tools, maintaining their integrity and reducing the manual management of these data.

Data migration has been performed with Safe FME program (Safe Software, 2018) mainly because it is user-friendly and its technical support makes this one of the most commonly used and extended extract, transform, and load (ETL) software worldwide. This ETL software allows you to reformat and connect different sources of information, maximising the uptime of DMI processes.

Figure 2.6 shows the FME canvas, where information from SITE (Smallworld) is migrated and integrated into the HYDORGIS tools.

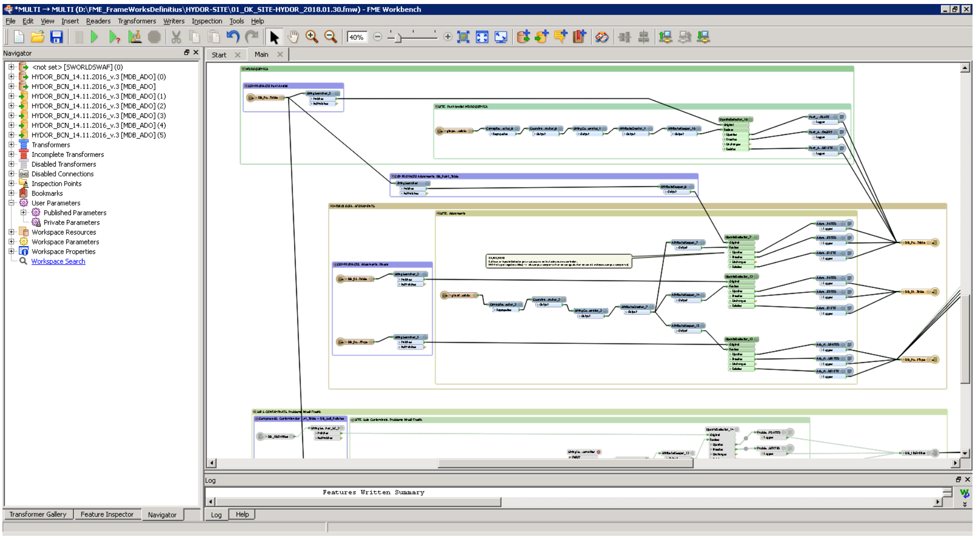


Fig. 2.6 FME canvas to transform information from SITE (Smallworld) to the HYDORGIS tools.

Figure 2.7 graphically describes our methodology for checking the existing data in HYDOR and the data stored in the SIEC/SICA systems that are not yet in HYDOR within an interval (or that have been modified in the SIEC/SICA systems; for example, modifications of piezometric levels). The variables that end with "_INI" and "_FI" control the source of data from SIEC and SICA systems. The variable that ends with "_MDB" controls data already found in the HYDOR database until the last date is stored. Therefore, to ensure a correct check, variables "_INI" and "_MDB" must be the same. Additionally, temporal variables "HG_DATA_INI", "HG_DATA_FI" and "HG_DATA_INI_MDB" are used to control the trickled migration of piezometric levels. Regarding the hydrochemical data, these variables are called "HQ_DATA_INI", "HQ_DATA_FI" and "HQ_DATA_INI_MDB".

The relationships among fields are crossed transferred; in other words, information stored in a single table of the BCASA systems will go to different tables in HYDOR. For accurate DMI processes, the following tasks are performed:

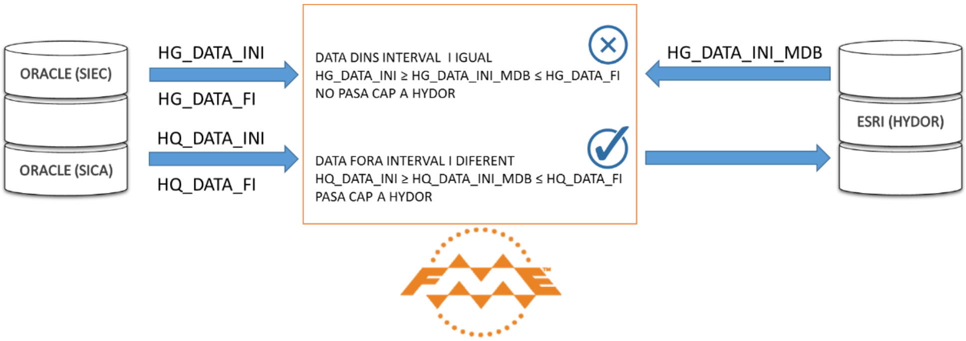


Fig. 2.7 Scheme of how our methodology checks existing data in HYDOR and data stored in the SIEC/SICA systems that are not yet in HYDOR (or that have been modified in the SIEC/SICA systems) within an interval.

- Check the validity of some data (*e.g.*, well depths, Z coordinates)
- Check possible mismatch of data (*e.g.*, duplication of data)
- Check and create important fields that are not in the original version of HYDOR

The last task can be performed thanks to the dynamic structure of the HYDOR database. It can be customised for different environments without losing its optimisation and interoperability among HYDORGIS tools with multiple sources of information.

Regarding hydrogeological data, the matching key fields in SITE and SIEC/SICA are "COD_SIS" and "COD_UBI", while the relationship of key fields for hydrochemical data in SITE and SIEC is only the "CODE" field.

In the SIEC and SICA systems, the field "NUM_X" (where X is "PIE" for piezometers or "EST" for station) is a key field in many tables related to SITE, indicating the variations in the installed sensors (repaired, new sensor, variation in its situation, etc.).

In the tables related to the instrumentation installed in each well, the field "IDE_X" (where X is "BOM" for pumping, "BMB" for pumping rates, "PIE" for piezometers or "AFO" for capacity) is a relevant field that indicates the position (the measurement point) of the device. “COD_UBI” store the mane of a zone where measures have been collected or are collected. For instance, the register "DJOM" stored in the “COD_UBI” field is the ID of a remote station where water level measurements are collected. Then, these datasets are sent automatically for their storage and analysis in the BCASA systems. Hence, "DJOM" remote station collects data from a well and a piezometer with "IDE_PIE" 15 and 16, respectively. Each IDE_PIE belongs to piezometers whose sensors (NUM_PIE) are connected to the same remote station "DJOM". The water level values of each IDE_PIE are stored in the SIEC table (BDE) BDE_NIV_PIEZO_ACT.

The DMI from SITE to HYDOR will be completed every time HYDOR is updated, and it has to follow the following three steps: (1) check SITE data already stored in HYDOR; (2) execute DMI processes from SITE to HYDOR; (3) compress HYDOR using ArcGIS (ESRI) tools.

The DMI processes from SIEC and SICA to HYDOR are incremental and are performed by entering two flags in all information loaded into the HYDOR tables for data tracking: (1) a BCASA flag and (2) a time flag. These flags are used to control the last date loaded into HYDOR from the BCASA systems. Hence, when a new interval of the DMI process is performed using date flags, the frameworks will check data with only these two flags, reducing the processing time.

There are two options with which to achieve the connection among these systems: (1) create new fields in HYDOR identical to the key fields in SITE (therefore, create fields in HYDOR identical to the key fields in SIEC and SICA) or (2) concatenate the SITE key fields (and therefore do the same in SIEC and SICA) and add them to HYDOR as a single key field. The last option seems suitable because it avoids modifications into internal structures of both the source and destination systems. Hence, the new concatenated key field will be stored with a unique ID in HYDOR tables. The concatenation is as follows:

unique ID in HYDOR = COD_SIS + "_" + IDE_PIE + "_" + COD_UBI + "_" + (flag of each original table from BCASA systems)

The flags of each original table are piezo (piezometer tables), pou (well tables) or panal (hydrochemical tables).

The key fields of the BCASA systems (*i.e.*, CODE, COD_UBI, COD_SIS, and IDE_PIE) are added to each of the tables corresponding to HYDOR.

Execution

After building the DMI model and creating DMI processes with FME during the execution of the frameworks, they cost more in terms of computational time. These frameworks (see Figure 2.6 as an example of one of these frameworks) are called from a batch file to streamline the process. This batch file is attached in Appendix B.

Validation and documentation

The validation of DMI processes is performed by comparing a well-known number of registers. This group of registers and taxonomies must be expanded to check all the sources of information and their natures. For instance, data can be loaded into HYDORGIS for a short period of time (*e.g.*, three years of measurements, including the data of two piezometers, with their head levels and hydrochemical analyses); then, the number of points to be accounted in the validation can be increased. In this stage, in addition, further errors in the DMI processes or in the origin systems can be detected; such errors include deduplication of registers in the origin systems and mistakes during data loading into the origin systems (*e.g.*, modifications in laboratory templates that automatically load hydrochemical data from the laboratory server).

The application of the proposed methodology in the BCASA shows that the proposed methodology is simple to implement, helps to check data quality from original sources of information, improves collaboration among departments involved in groundwater management and facilitates common hydrogeological analyses.

Example of application: the Alternative Water Plan of Barcelona

The information migrated and integrated from civil works, scientific field campaigns and BCASA systems to HYDORGIS tools are applied in the Alternative Water

Plan (PLARHAB). This plan has been suggested to the Barcelona City Council every five years since 1998 and aims at improving groundwater uses and applications in the city.

In general, the geology of Barcelona can be simplified into four zones according to its hydrogeological characteristics: the Collserola Mountains (mainly metamorphic materials), the Llobregat and Besòs deltas (alluvial fans) and the Barcelona plain (alluvial fans) (Vázquez-Suñé et al., 2010). The hydraulic flow is dependent on not only the water balance but also the anthropic impacts: underground infrastructures (tunnels or car parks) and dewatering that are distributed throughout the city. The hydraulic gradient, in general, from the mountains to the sea, crosses different geological units and is one of the conditions that give the characteristic composition of groundwater bodies. In addition to the main natural recharge sources (mountains and the Besòs River), there are some anthropic sources (sewage and water supply networks, mainly) that artificially recharge the aquifers, directly impacting the groundwater quality.

The Collserola Mountains and the topographic highs of the alluvial plains of the Barcelona plain are areas where rainwater is captured by wells and springs. We do not have much information about this zone because there are close to the mountains (less permeable, in general) so here are not many control points,. Therefore, these areas are not relevant for their use. In general, the groundwater composition at the sampled points is poor in most ions; due to the fact that they cross materials with a low content of magnesium and calcium (see Figure 2.8).

Notably, in the middle area of the Barcelona plain, the concentration of ions in the groundwater is less than those of the Llobregat and Besòs Rivers (see Figure 8). Conversely, the composition of waters sampled from points close to the Besòs River is very similar to the composition of the river water. This may indicate that (i) materials at each point in that zone are very permeable and (ii) a significant contribution of aquifer water comes from the river. Additionally, this implies that there is a contribution of pollutants that moves from the river to the aquifer; once added to the flow of groundwater, their concentration is reduced because of the ability of the aquifers to reduce their concentrations by redox processes and because of water mixing from the different sources of recharge (Jurado et al., 2013; Vázquez-Suñé et al., 2016b).

Figure 2.9 shows the evolution of representative points in each city area: topographic highs of the Barcelona plain (SALL), middle area of the Barcelona plain

(DANO), around the Besòs River (PUIG), close to the Llobregat River (TENB) and near the sea (WEPO). However, the SALL point, despite being close to the rain recharge zone (with a low concentration of salt), has higher electrical conductivity than those of the DANO point, located in the middle of the city. The highest values of electrical conductivity at the SALL point can be a result of an additional contribution of salts, probably due to losses in the wastewater network and/or losses in the drinking water network.

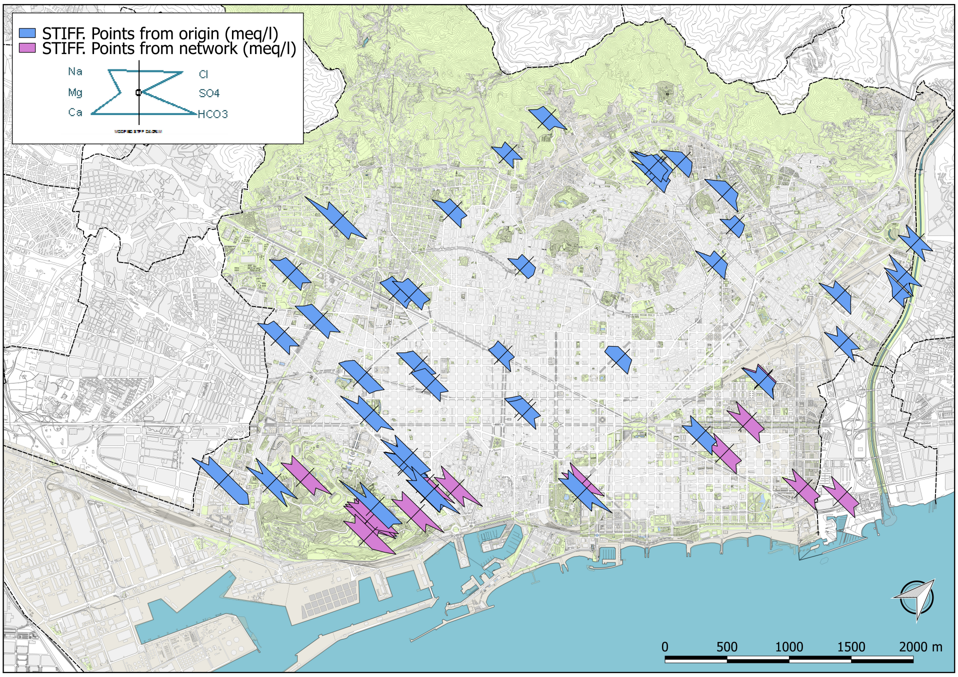


Fig. 2.8 Stiff diagrams from winter 2017. Diagrams in blue represent samples obtained from wells and piezometers, while are samples obtained at points in the groundwater supply network are shown in pink.

Figure 2.10 graphically describes the spatial distribution of the electrical conductivity. Most of the sampled points have values less than 2,000 μS/cm, with the exception in the area close to the Besòs River and the sea (Poble Nou area). Saline intrusion in the Poble Nou zone has been increasing because of dewatering performed during the construction of new infrastructures. This area is greatly affected by marine intrusion and has been controlled by a detailed monitoring network installed during the construction of one of the largest civil works projects conducted in the

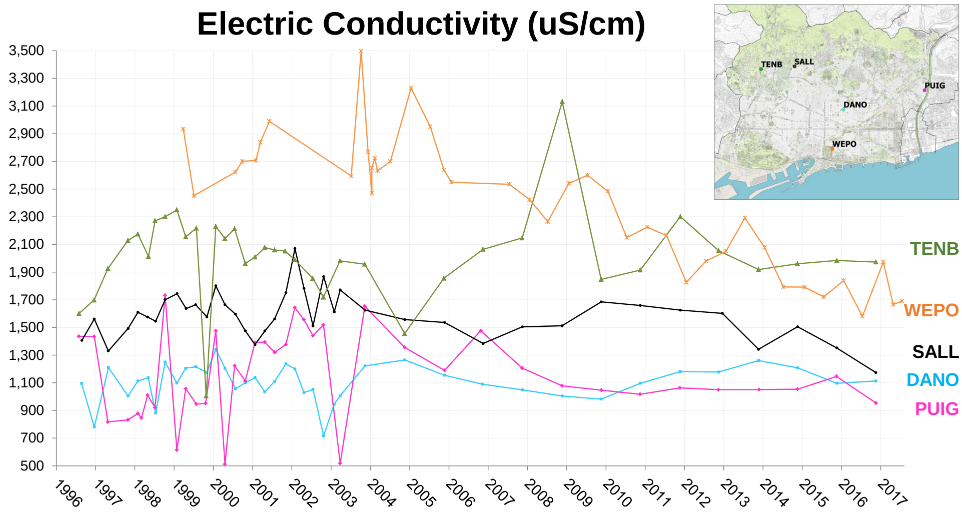


Fig. 2.9 Temporary evolution of electrical conductivity (μS/cm) of representative points in the upper part of the Barcelona plain (SALL), middle area of the city (DANO), surroundings of the Besòs River (PUIG), zone close to the Llobregat River (TENB) and area close the sea (WEPO). Their positions are schematized on the right.

city. According to field campaigns, most of the points around the Poble Nou area have values greater than 3,000 μS/cm, reaching values that exceed 55,000 μS/cm at points measured in the upper aquifer. Some measures will be taken to control the seawater intrusion in the zone.

Concentrations of nitrates, sulphates and other pollutants in groundwater may indicate losses from the wastewater and leakage from industrial sites because the measured concentrations are often higher than the concentrations present in other sources of recharge (Vázquez-Suñé et al., 2016b).

The general composition of sampled waters may indicate that use of this water is appropriate for irrigation (despite having a risk of soil salinisation due to high electrical conductivity) and cleaning of city infrastructures (streets and rainwater retention tanks).

The city of Barcelona has already implemented alternative uses (mainly irrigation) of groundwater for many years. The results obtained show the increasing quality of groundwater over the years, leading to reduced treatment costs before its use. For another use (such as drinking water), a more detailed hydrochemical analysis of non-major elements (such as heavy metals) and emerging contaminants (such as pesticides and drugs) should be performed throughout the city to continue improving groundwater use by designing the most suitable treatment.

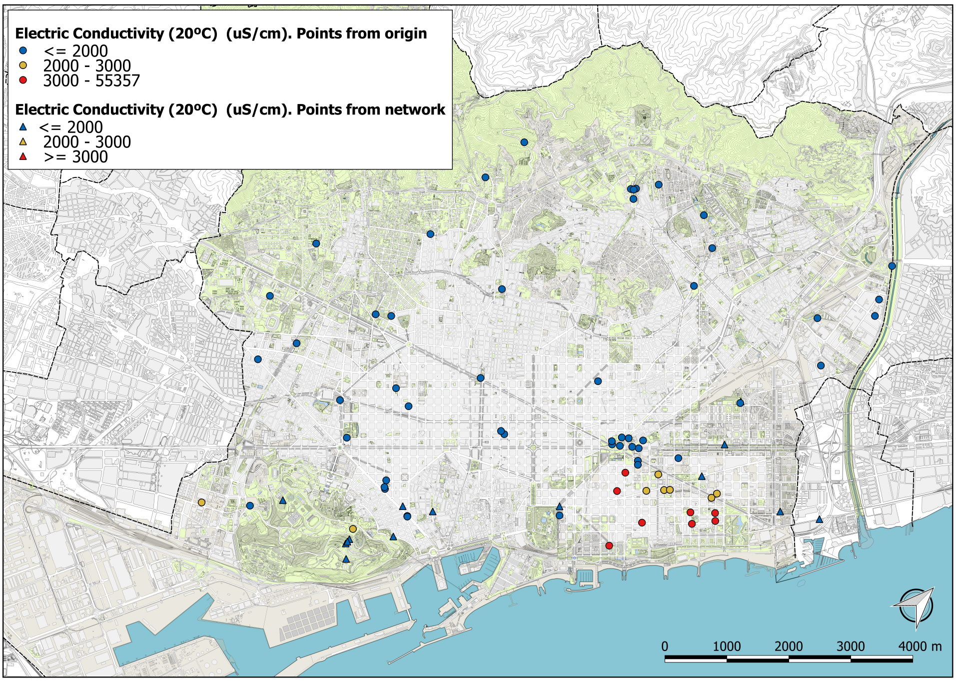


Fig. 2.10 Electric conductivity map. Winter campaign (2017). Dots represent samples collected directly from wells and piezometers, while samples obtained at points in the groundwater supply network are represented by triangles.

2.4 Conclusions

To develop a data infrastructure or to connect it with information communication technology (ICT) tools, data migration and integration (DMI) of all sources available is a big step forward for geodata governance when multiple sources of information have to be homogenised with minor interruptions during production to deliver quality information to users and help them in further analyses with a unique platform.

In this manuscript, we presented a novel methodology for facilitating and improving data migration and integration of multiple data infrastructures to new systems or to connect them with ICT tools. This new methodology allows higher performance of all data collected, ensuring proper groundwater management by improving data governance.

The application of this methodology by the Barcelona City Council, using Barcelona water cycle (BCASA) systems, optimises the analysis process and shares the results of their hydrogeological assessments (*e.g.*, the automatic sensor maintenance and groundwater quality control). The ICT tools implemented in the BCASA systems facilitate the manipulation, visualisation and analysis of the hydrogeological data in a unique platform (HYDORGIS tools), while the original data infrastructures continue storing the groundwater information. The DMI models performed in BCASA systems can be easily adapted to other external datasets, increasing the volume of quality data to improve the understanding of the groundwater system behaviour and improving the monitoring network in the city. Other benefits observed during the implementation of this method are that the DMI methodology aids in checking data quality in original sources of information and encourages collaboration among people with different skills from different departments.

The methodology proposed here can be implemented in any kind of DMI process, to develop data warehouses, spatial data infrastructures or to implement ICT tools for further analyses, thus improving data governance.

Chapter 3

Hydrogeological analysis. HYYH and MJ-Pumpit

Summary

The quantification of the hydraulic parameters is important to support decision making in environmental impact assessment, water resources evaluation or groundwater contamination remediation, among others. These kind of parameters derived from aquifer tests usually encompasses a vast amount of data (spatial and non-spatial) for management and analysis. To achieve this in a clear and understandable manner, the GIS environment is a useful instrument. Development of innovative software to analyse pumping tests in a GIS platform environment to support the hydraulic parameterization of groundwater flow and transport models is presented in this paper. This new platform provides three interconnected modules to improve (a) pumping test interpretation code through a user-friendly interface, (b) pumping test data visualisation supported by a set of tools that perform spatiotemporal queries in a GIS environment and (c) the storage and management of hydrogeological informa-

Based on: Criollo, R., Velasco, V., Vázquez-Suñé, E., Serrano-Juan, A., Alcaraz, M., García-Gil, A. An integrated GIS-based tool for aquifer test analysis. Environmental Earth Sciences (2016) 75: 391. doi:10.1007/s12665-016-5292-3.

tion. Additionally, within the GIS platform, it is possible to process the hydraulic parameters obtained from the pumping test and to create spatial distribution maps, perform geostatistical analysis and export the information to an external software platform. Finally, a real-world application in the area of Barcelona (Spain) has shown the usefulness of the tools developed in support of hydrogeological analysis.

3.1 Introduction

The quantification of hydraulic parameters such as transmissivity (T), hydraulic conductivity (K), storativity (S), and specific storage (Ss) is important for hydrogeological assessments and for the development of groundwater flow and transport models to support decision making in environmental impact assessment, groundwater contamination remediation, water resources evaluation or site monitoring (Rogiers et al., 2012).

There are several methods to determine hydraulic parameters. The selection of a suitable method depends on the purpose of the research and the required degree of accuracy (Vukovic and Soro, 1992). The pumping test is the most commonly used method to obtain aquifer parameters and generally leads to reliable hydraulic parameters. The reliability of the results depends on many factors such as the following: (1) availability of wells and observation points for testing, (2) quality of the pumping test and the subsequent management and analysis of the data obtained and (3) accurate information of aquifer geometry and hydraulic boundaries (Cheong et al., 2008). Furthermore, a number of specific issues must be considered when dealing with aquifer tests. On the one hand, each problem requires specific solutions (different hydrogeological conditions, different results, different study area and wellbore characteristics, etc.). On the other hand, regarding aquifer test interpretation, a selection of the proper graphical, analytical and numerical solutions for the interpretation should be considered. Moreover, the use of full or partial datasets for the analysis should be taken into account. In this sense, the use of comprehensive tools to store, manage and visualise the vast amount of available data becomes a necessity to focus the analysis and leads to an accurate interpretation.

Therefore, it is not surprising that, currently, there is a great amount of software oriented to manipulate and facilitate the calculations of hydraulic parameters from aquifer tests. These computer tools reduce the calculation time and allow a better and more accurate determination of hydraulic parameters. For instance, WTAQ

(Barlow and Moench, 1999), WIGAEM (Bakker, 2009) or WELLS (Vesselinov et al., 2009) apply analytical solutions for pumping test analysis, but there is still a lack of tools for the pre- and post-processing of input and output data. This is an important problem to be solved because the size of the datasets to be processed is continuously increasing. This growth is closely linked to the importance of water resources, which is becoming more crucial for human communities and because of the greater complexities of the regional groundwater numerical models being used currently. In addition, the hydraulic data available for integration into groundwater numerical models usually has a very diverse origin and format and, therefore, a chance of bias in the interpretations. Consequently, it becomes necessary to have effective instruments that facilitate the pre-process, the visualisation, the analysis and the validation (*e.g.* graphical analysis techniques) of this great amount of data.

Software platforms such as MATLAB (The MathWorks) or MS Excel (Microsoft) seem to be more appropriate environments for pre- and post-processing hydrogeological data because of their rich set of tools that are oriented to analyse and visualise data results. Additionally, these software platforms provide a graphical user interface (GUI) for developing specific tools. Some applications oriented to aquifer test data analysis have been developed in the MATLAB environment such as HYTOOL (Renard, 2008) or CHOW (Zhan et al., 2001), which provide a user-friendly interaction with the scripts for aquifer test analysis. In the same way, MS Excel is a spreadsheet program that allows analysis and plotting of input and output data by using its built-in functions and its Visual Basic for Application (hereinafter VBA) macro option. Some examples of specific tools for pumping test data analysis developed in this context are USGS Spreadsheets (Halford and Kunianksy, 2002), Molano (2013) or Johnson and Cosgrove (2001). In addition, the latter can make draw down contours and mass balance graphs of RADFLOW code.

Computer programs that incorporate additional methodologies to process visualise and interpret aquifer test data give one step beyond. Some examples are AQTESTSOLV (Duffield, 1989), Well32 (GeoSoft, 1993), AquiferWin32 (Rumbaugh and Rumbaugh, 1997), Aquifer Test (Röhrich, 2002) or MLU (Hemker and Post, 2010). While these software contains many powerful features for performing aquifer tests analysis, further developments in the management of large datasets provided by the aquifer tests and by its interpretations would significantly improve the resulting hydraulic parameterization of the study area. Besides an appropriate storage of all available data and documentation of the procedures used for the interpretation of the pumping tests, it may facilitate the further updating of the initial parameterization of the hydrogeological model and may also guarantee its future reuse by third parties for different objectives. In this regard, additional advances have been achieved by software such as PIBE (Diputación de Alicante, 2006) or Ephebo (UPC, 2002).

Despite of these advances, further analysis should be focused on improving the interpretation and validation of the hydraulic parameterization provided by aquifer tests such as the cross-analysis with other datasets (*e.g.*, aquifer geometry or hydraulic boundary conditions). In this sense, the use of a Geographical Information System (GIS) environments represent an optimal solutions for the integration of the aquifer tests data with other relevant datasets (*e.g.*, geological or meteorological datasets) and enable the straightforward queries, search and retrieval of portions of information provided by different sources. Common GIS applications for groundwater research include tools that generate spatiotemporal queries of different hydrogeological parameters. ArcHydro (Maidment, 2002; Strassberg, 2005), CUAHSI (Maidment, 2005) or GMS (Jones et al., 2004) are examples of these types of GIS-based tools, which take advantage of the GIS platforms to manipulate and visualise aquifer information, giving another dimension to the analysis process.

By emphasising pumping tests and decision making GIS-based tools, Rios (2011) developed software (uWATER-PA) for non-specialised issues of groundwater pumping impacts limited by one specific analytical solution for calculations.

The foregoing improvements to analyse, visualise and interpret aquifer tests constitute a relevant advantage for the estimation of hydraulic parameters over traditional analysis. However, additional refinements are still needed such as storing, managing and analysing all the available data provided by the aquifer tests interpreted by different methods into a GIS environment. Indeed, the main objective of this work is to provide these needed refinements. To perform that, we provide a package of tools for collecting, managing, analysing, processing and interpreting data derived from pumping tests in a GIS environment. In this sense, the tools developed allow us to apply different methods to interpret aquifer test data in several scenarios taking into account all the information available related with the study zone. Using the standardised spatial database aids to a proper data management. This will increase the understanding and the knowledge in the study area and, thus, will help the development of future projects.

This chapter is distributed as follows: design and features of the developed software are briefly described in section two. In section three, an application of these tools in a study area is presented and, finally, the conclusions are presented together with additional discussion on the advantages and disadvantages of the software.

3.2 Pumping test software platform

To facilitate the management and interpretation of pumping tests in a GIS environment, four main requirements should be considered: (i) storing and managing hydrogeological data related to pumping tests in an uniform structure, (ii) pumping test data processing and visualisation in a GIS environment, (iii) specific tools to interpret or re-interpret pumping tests considering different methods and (iv) interoperability with external software for further analysis to complete hydrogeological studies. These requirements mentioned have been reached by the methodology used in the next section.

Software design

To develop the software platform presented in this paper, the following technical criteria were used as guidelines:

- A geospatial database to integrate the sets of data from pumping tests. The developed database requires tools and methodologies to simplify the end-user's tasks (user interface and protocols for data exchange).
- GIS environment with the following characteristics:
    - Suite of inherent tools in a GIS environment (*e.g.*, mapping, georeferencing, geostatistical tools, etc.).
    - Specialised tools to analyse and integrate hydrogeological data (*e.g.*, to consult points involved in the test, pumping test duration, parameter values estimated, etc.) fed by a structured database.
- Interoperability with external software for further analysis such as aquifer test analysis.
- Statistics tools, filters and queries, and different methodologies to solve different aquifer scenarios to guarantee an optimal interpretation.
- Post-processing tools. Once the interpretation of pumping tests with external software is performed, the results should be imported again into the GIS

platform to continue with the spatial analysis (*e.g.*, useful to prepare maps of parameters to export them to a modelling package). In addition, this analysis can be supported by accurate information on aquifer geometry, hydraulic boundaries (Cheong et al., 2008) or other spatiotemporal data in a clear and understandable manner.

Software features

The presented work forms part of a wider on-going framework to facilitate detailed hydrogeological modelling that includes additional hydrochemical and geological GIS-based analysis tools and a geospatial database termed HYDOR. These tools are described in Velasco (2013) and Velasco et al. (2012a; 2012b; 2014).

The innovative set of analysis tools and methodologies oriented to analyse, visualise and interpret aquifer test data are explained above. This set of tools was developed as a set of modules, which are interconnected to each other and simplifies the end-user's tasks through the integration of: (1) a module for the analysis of pumping tests (MJ-Pumpit), (2) a geospatial database to record and manage data involved in pumping tests (HYDOR) and (3) tools for advanced spatiotemporal analysis built in the same GIS environment (HYYH toolbar) (see Figure 3.1). The module integration is possible thanks to the compatibility among the software platforms used in to build each module. An optimal workflow is possible preserving a complete data control with the geodatabase.

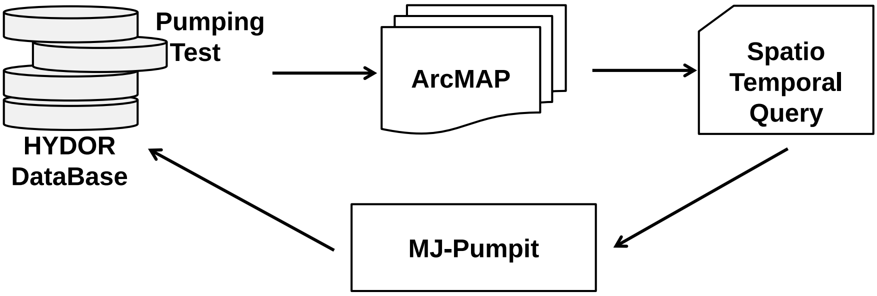


Fig. 3.1 Software architecture. Data for the spatiotemporal query and for MJ-Pumpit are managed by the spatial database and visualised in the GIS platform.

MJ-Pumpit

MJ-Pumpit is a GUI for MariaJ code (Carbonell et al., 1997), hereinafter MJ, which simplifies the data handling and model selection needed to execute MJ code. Thus, MJ-Pumpit has been developed as a MS Excel Add-in, programmed in VBA, following the methodology proposed by Serrano Juan (2016). The use of MS Excel spreadsheets as a platform is based on the following reasons: (i) it is a highly stable software; (ii) it is a widely used software; (iii) it ensures a portable and easily handled format; (iv) it allows an easy performance of data queries (*e.g.*, to check potential mistakes), analyses and plots; (v) it has a user-friendly interface to create numerical and statistical computations, among others. In this way, MJ-Pumpit has the available MS Excel tools to analyse and plot data to ease the pumping test interpretation with analytical models for different aquifer scenarios included in MJ code.

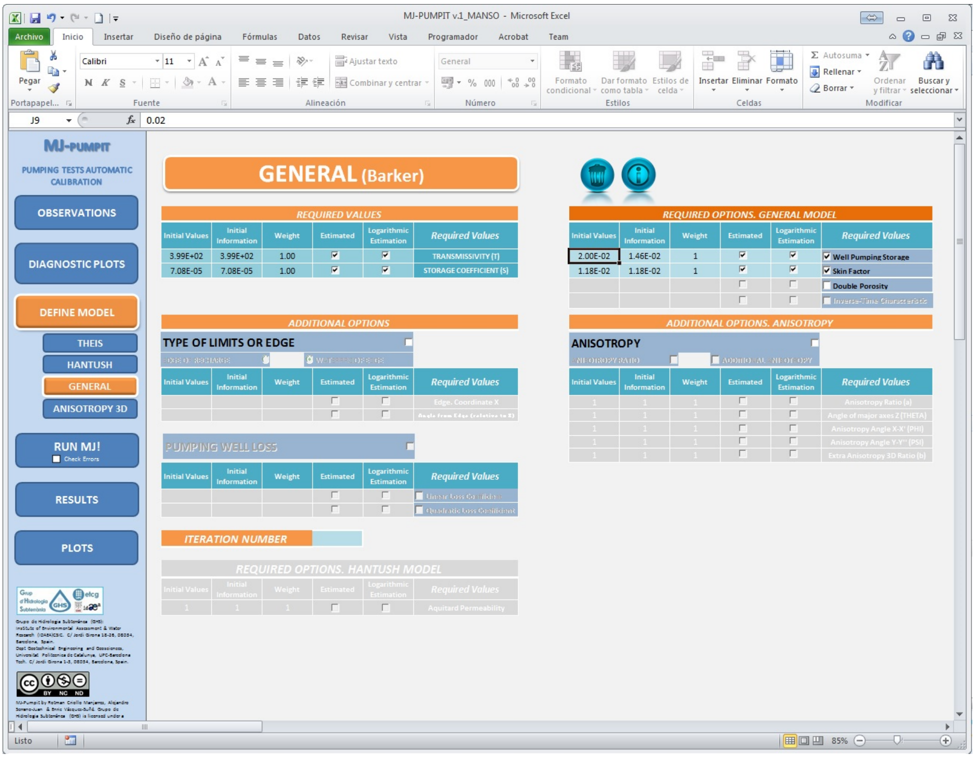
The MJ code (FORTRAN-based) is a software for automatic calibration of pumping test, which includes several analytical models to interpret the test. It obtains the best aquifer parameters for the considered model and fits the results obtained with the observations data by minimising the objective function. Input data, statistical analysis and information about the iteration process are printed in the output file, which is also uploaded into MJ-Pumpit and displayed automatically.

MJ-Pumpit allows the user to load pumping test data by using its own query interface. Additionally, a command developed directly in the GIS platform enables retrieving the information for the selected wells and time intervals by exporting data into MJ-Pumpit, adding a previous spatio-temporal analysis.

MJ-Pumpit features

The GUI of MJ-Pumpit is composed of different commands that facilitate the interpretation process of pumping tests and allows the following:

- Interchange data with HYDOR database through the GIS platform. Export/import the pumping test data and its interpretation from or to MJ-Pumpit.
- Error typing avoidance during the field data insertion process.
- Field data quality analysis improvement in MS Excel before executing the code.
- Ensuring an appropriate selection of the conceptual model using diagnostic plots.



fx 0.02

GENERAL (Barker)

REQUIRED VALUES

Initial Values	Initial Information	Weight	Estimated	Logarithmic Estimation	Required Values
3.99E+02	3.99E+02	1.00	☑	☑	TRANSMISSIVITY (T)
7.08E-05	7.08E-05	1.00	☑	☑	STORAGE COEFFICIENT (S)

ADDITIONAL OPTIONS

TYPE OF LIMITS OR EDGE ☐

○ EDGE OF RECHARGE ○ WATERPROOF EDGE

Initial Values	Initial Information	Weight	Estimated	Logarithmic Estimation	Required Values
			☐	☐	Edge. Coordinate X
			☐	☐	Angle from Edge (relative to X)

PUMPING WELL LOSS ☐

Initial Values	Initial Information	Weight	Estimated	Logarithmic Estimation	Required Values
			☐	☐	☐ Linear Loss Coefficient
			☐	☐	☐ Quadratic Loss Coefficient

ITERATION NUMBER

REQUIRED OPTIONS. HANTUSH MODEL

Initial Values	Initial Information	Weight	Estimated	Logarithmic Estimation	Required Values
1	1	1	☐	☐	Aquitard Permeability

REQUIRED OPTIONS. GENERAL MODEL

Initial Values	Initial Information	Weight	Estimated	Logarithmic Estimation	Required Values
2.00E-02	1.46E-02	1	☑	☑	☑ Well Pumping Storage
1.18E-02	1.18E-02	1	☑	☑	☑ Skin Factor
			☐	☐	☐ Double Porosity
			☐	☐	☐ Inverse-Time Characteristic

ADDITIONAL OPTIONS. ANISOTROPY

ANISOTROPY ☐

ANISOTROPY RATIO ☐ ☐ ADDITIONAL ANISOTROPY

Initial Values	Initial Information	Weight	Estimated	Logarithmic Estimation	Required Values
1	1	1	☐	☐	Anisotropy Ratio (a)
1	1	1	☐	☐	Angle of major axes Z (THETA)
1	1	1	☐	☐	Anisotropy Angle X-X' (PHI)
1	1	1	☐	☐	Anisotropy Angle Y-Y'' (PSI)
1	1	1	☐	☐	Extra Anisotropy 3D Ratio (b)

Fig. 3.2 MJ-Pumpit: the model selection spreadsheet introduces several models of aquifer environments (buttons on the left; in this case, the Barker (1988) model is selected and has different options available such as well pumping storage, skin factor, and anisotropy, among others).

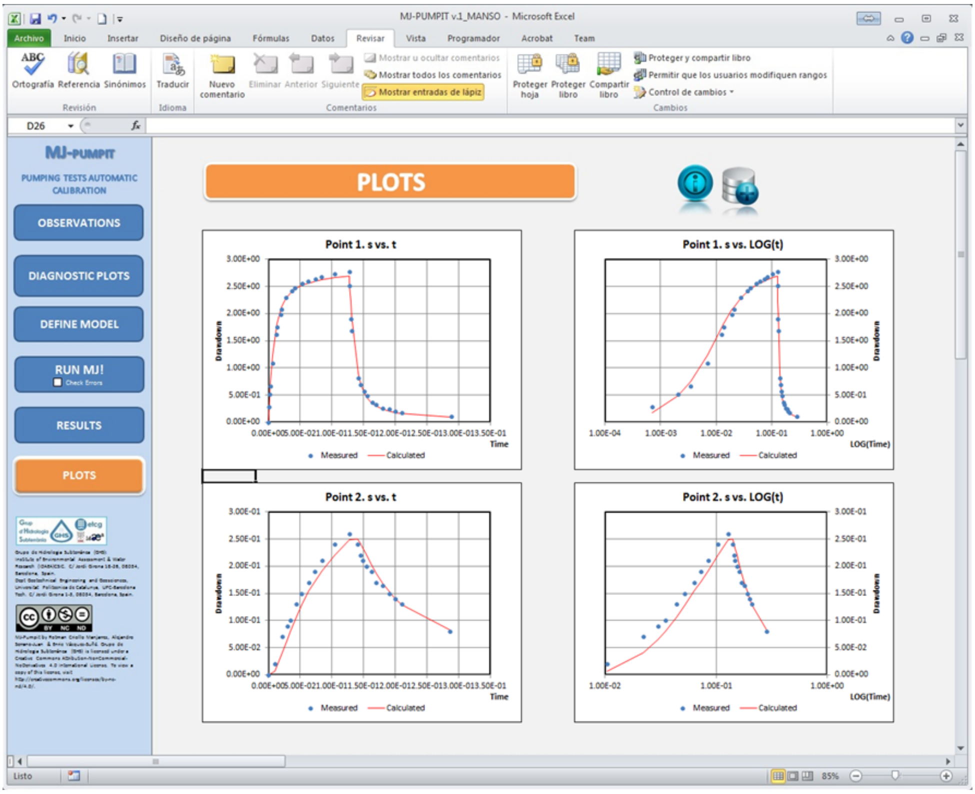


Fig. 3.3 MJ-Pumpit: results spreadsheet allows us to analyse the MJ iteration process (button *results*), plots of fitted values (button *plots*) and export parameter values to the HYDOR database for additional spatiotemporal analysis in a GIS environment (button with a database icon).

- The integration of different aquifer scenarios by different methodologies of pumping test analyses and other characteristics are available such as edges (recharge or waterproof), pumping well loss, double porosity, among others (see Figure 3.2).
- The results are displayed from the iterative process through automatically generated reports and graphical plots of fitted and measured values (Figure 3.3). Their formats can be changed per the user's requirements with MS Excel capabilities.
- MJ-Pumpit can be used in six different languages.

Spatial database

The spatial database HYDOR was developed by following the structure of a Personal Geodatabase (ArcGIS, ESRI) and it was created following international standards and taking into account other hydrogeological databases. Its relational structure aims to store, operate and relate several topics in different datasets to manage geological, hydrochemical, hydrogeological, hydrological, geophysical and meteorological information. In addition, to standardise and harmonise the information collected, the HYDOR geodatabase allows insert data in different formats following a protocol that was developed to import a large volume of data at the same time or "one by one". For further information about HYDOR see Velasco et al., (2012a; 2012b) and Velasco (2013).

Although the data model of the hydrogeological database described here was implemented within ArcGIS, most of these concepts are sufficiently flexible to enable implementation in other platforms (Velasco et al., 2014).

Developed in the same way, regarding the information on pumping tests, a group of tables have been created to support the storage and management of its data. In this sense, to facilitate data handling for any user, the pumping test information collected is separated into different datasets:

- Geographic features stored in *DB_Points* table (*e.g.*, point names, coordinates, source of information, etc.).
- Specific characteristics of wells such as radius, depth or screens are collected in *DB_ Wells* table.

- Test characteristics and observations are included in *DB _Pumping TestSpecifications* and *DB_Pumping TestObservations* tables (*e.g.*, flow and drawdown timings, cumulative volumes, pumping well, etc.).
- Pump characteristics are collected in *DB_Pumps*.
- Initial interpretations or interpretations made with MJ-Pumpit are located in *DB _Pumping TestInterpretations*.

Additionally, to ensure the standardisation of data recorded in the database, different topics are arranged into list codes such as test types, units of measure or relevant data (*e.g.*, type of pumping test), and additional details of the test, among others, are compiled. A simplified entity relationship diagram for the database structure is presented in Figure 3.4 and shows the interaction between the measurements, pump characteristics and pumping test interpretations. To ensure the right database functionality, the environment has mechanisms to facilitate data transcription and to avoid its inconsistency during their insertion. Other specialised tools to simplify the spatial and temporal queries of pumping test data will be discussed below.

Spatiotemporal query tools

The spatiotemporal tools developed in the present work are an extension of the HYYH toolbar, which was designed to analyse and visualise different hydrogeological measurements and the results of field tests (including pumping tests) stored in the HYDOR database for a complete hydrogeological analysis. The different tools integrated in the HYYH toolbar are shown in Figure 3.5 and fully described in the following sections. These tools were created with ArcObject, a developer kit based on the Component Object Model (COM) by using the Visual Studio (2010). These platforms were selected because: (i) it allow reuse and use external objects as their own, (ii) its large set of standards for implementing and using objects and for inter-object communication, (iii) COM offer interact across process and computer boundaries as a easily manner within a single process, among others (Microsoft, 2018).

Hydrogeological data analysis tools

This set of tools makes queries regarding the evolution of the piezometric levels and the control of the aquifer abstractions (see Figure 3.5). This information can be analysed individually or jointly with other information stored in the database. The ***"temporal query"*** can discretise the evolution of the desired parameter by

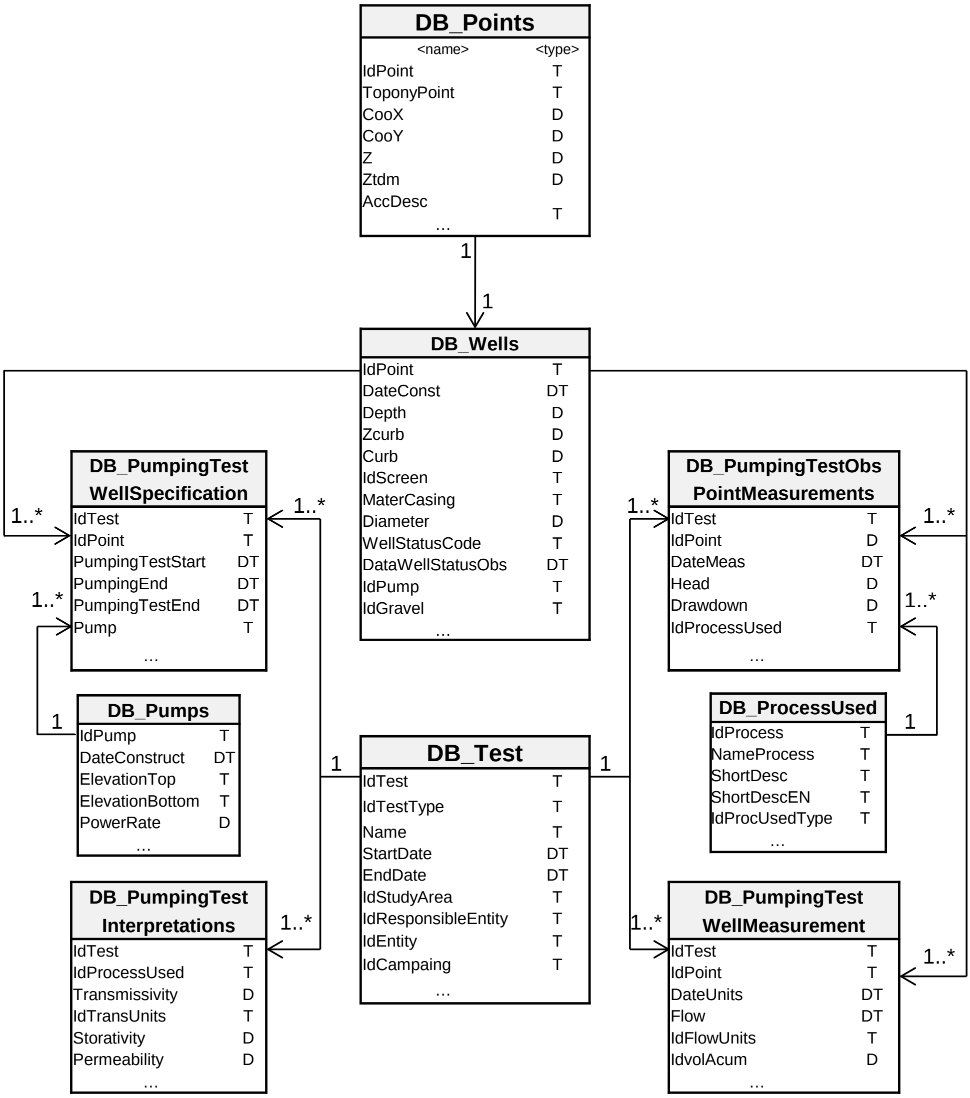
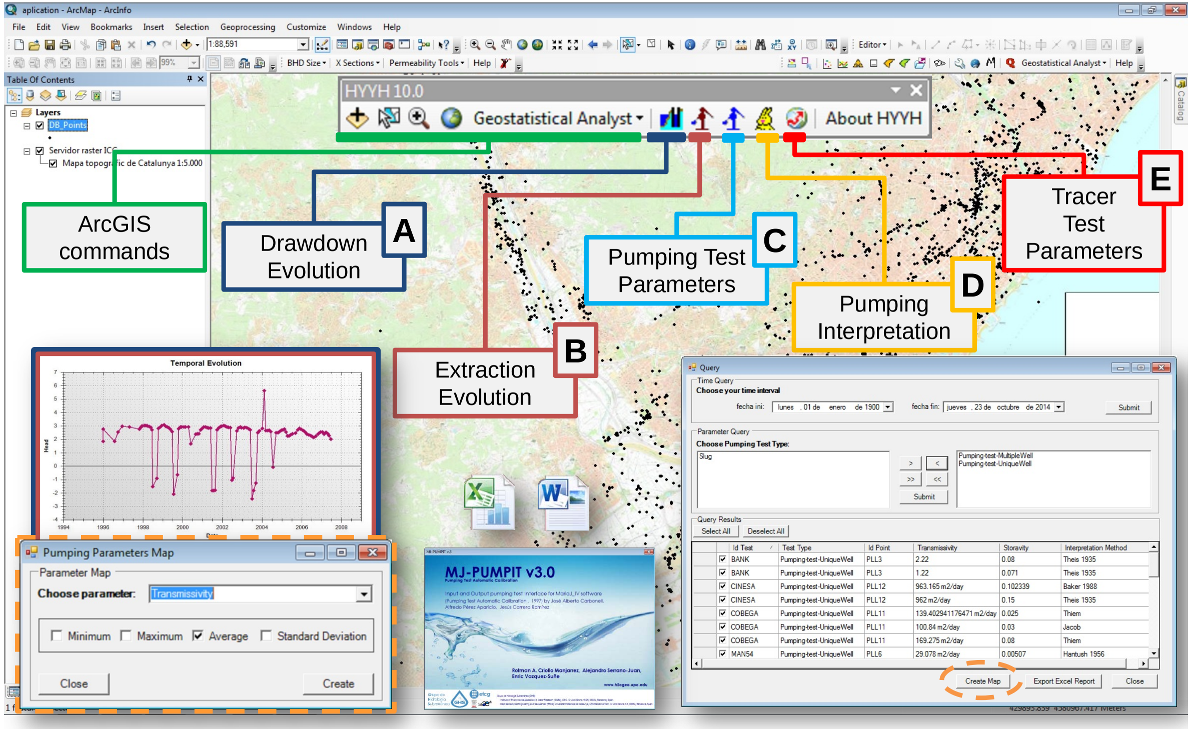


Fig. 3.4 The conceptual diagram represents the main fields that comprise the hydrogeological database referred to in the pumping test. The 1 and 1* represent the cardinality of the relationship between tables. T, D and DT are text, double and date time formats types, respectively.

displaying the information such as campaign name, measurement date or measure point in a table. For further analysis, the user can display the temporal evolution of the data selected (***"Graph"*** command), export data to MS Word (***"RTF Report command"***), export data to MS Excel (***"Excel Report command"***), or use the ***"Create Map"*** command. The functionality of the latter command is used to calculate basic statistical data (minimum, maximum, average and standard deviation), after to be selected the required points in the screen, for each selected parameter for a given period of time. These values are displayed in a shapefile, which can be used for future geostatistical analysis through the geostatistical analyst tools inherent in ArcGIS without leaving the platform.



Pumping Parameters Map

Parameter Map

Choose parameter: Transmissivity

☐ Minimum ☐ Maximum ☑ Average ☐ Standard Deviation

Close Create

Query

Time Query

Choose your time interval

fecha ini: lunes , 01 de enero de 1900

fecha fin: jueves , 23 de octubre de 2014

Parameter Query

Choose Pumping Test Type:

Slug

Pumping-test-MultipleWell
Pumping-test-UniqueWell

Query Results

	Id Test	Test Type	Id Point	Transmissivity	Storavity	Interpretation Method
☑	BANK	Pumping-test-UniqueWell	PLL3	2.22	0.08	Theis 1935
☑	BANK	Pumping-test-UniqueWell	PLL3	1.22	0.071	Theis 1935
☑	CINESA	Pumping-test-UniqueWell	PLL12	963.165 m2/day	0.102339	Baker 1988
☑	CINESA	Pumping-test-UniqueWell	PLL12	962 m2/day	0.15	Theis 1935
☑	COBEGA	Pumping-test-UniqueWell	PLL11	139.402941176471 m2/day	0.025	Thiem
☑	COBEGA	Pumping-test-UniqueWell	PLL11	100.84 m2/day	0.03	Jacob
☑	COBEGA	Pumping-test-UniqueWell	PLL11	169.275 m2/day	0.08	Thiem
☑	MAN54	Pumping-test-UniqueWell	PLL6	29.078 m2/day	0.00507	Hantush 1956

Fig. 3.5 Scheme of HYYH toolbar tools available in ArcMAP. Hydrogeological data analysis tools (A and B tools) and hydrogeological test analysis tools (C, D and E tools) have specialised commands to export data selected for other platforms, temporal evolution plots, and created parameter maps.

Hydrogeological test analysis tools

As stated above, the selection of the pumping tests (see Figure 3.5) is performed by choosing a set of points on the screen that represents points involved in a given aquifer test. The analysis tools of the hydrogeological test (tracer and pumping test interpretations) allow to visualise the data stored in the database with the spatio-temporal query. Once the temporal discretisation query of the selected points in the ArcMap (ESRI) is performed, a list of data related to the hydrogeological test is given. This data, such as campaign name, date of test, parameter values obtained by different aquifer field tests, among others, can be exported to other platforms (MS Excel or MS Word) and make effective reports or create maps of statistical values of the selected parameter for additional uses (*e.g.*, hydraulic parameter maps to be exported to the numerical model, see Figure 3.6).

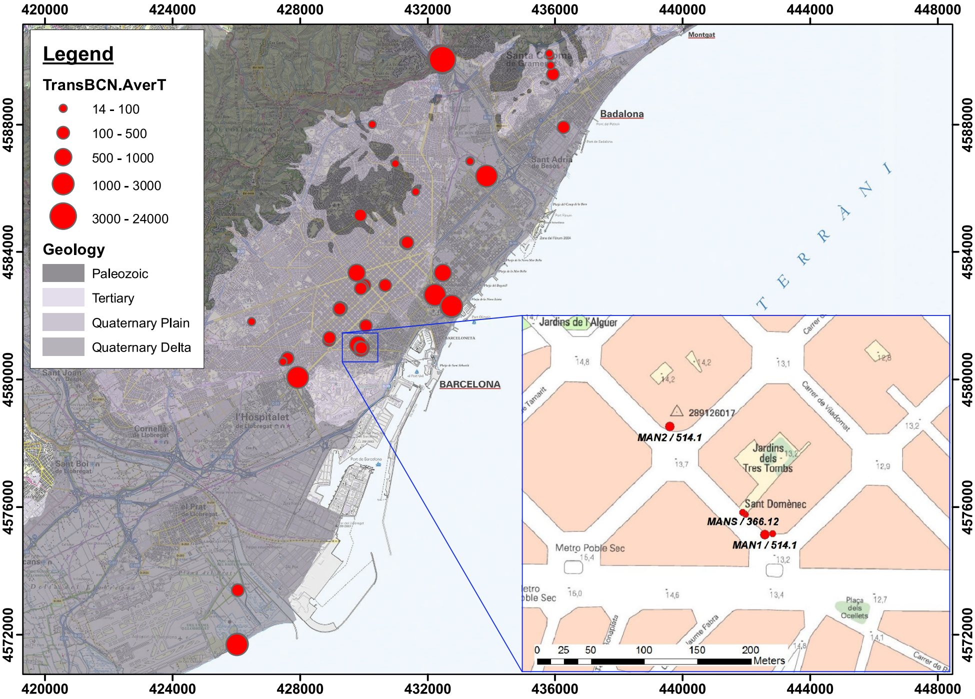


Fig. 3.6 Transmissivity map created using the hydrogeological test analysis tools. These types of maps can aid the modelling groundwater flow and solute transport. With geostatistical wizards available in ArcGIS, users can extrapolate this information and create surface maps for additional uses.

If there are one or more pumping tests without interpretations or with old interpretations in the study area, the user can make a new pumping test analysis simply and quickly.

Pumping test analysis tool

This command has been developed to analyse pumping test data involved during the test, which is composed of the following commands.

Querying the test data

Similar to the previous sections, the spatial and temporal discretisation of the study area gives a list of several available fields in the database related to pumping tests (*e.g.*, test name, interval time, or number of points involved in the test). For each test selected, the information can be exported to another platform to make reports (MS Word or MS Excel) or to interpret the test with MJ-Pumpit.

Link to MJ-Pumpit

This command allows the export/import of selected tests and related data for a new interpretation to or from the MJ-Pumpit GUI. Once a user obtains the hydraulic parameters with MJ-Pumpit, it is possible to load them back into the GIS platform (storing the analysis data and results in the geospatial database) and to reinforce the hydrogeological study by complementing it with other datasets for a complete hydrogeological analysis using other information treated with the tools described above (see Figure 3.7 as an example).

3.3 An example of application: the urban area of Barcelona (Spain)

The designed software has been applied in The Metropolitan area of Barcelona, which is located in northeastern Spain (Figure 3.7). Barcelona is limited by the Serra de Collserola (NNE), the Mediterranean Sea (SSW) and two rivers: Besòs (NE) and Llobregat (SW). The latter is delimited by Montjuïc hill next to the seashore. From the geological characteristics, different hydrogeological units can be identified: Paleozoic materials, made by shales and granites, are located in the high part of the city; Montjuïc hill slope made up of marls, sandstones and sands from Tertiary period; and Quaternary aquifers which can be separated in alluvial - deltaic sediments in topographic low areas and piedmont cones - coarse alluvial sediments located in the intermediate areas (Vázquez-Suñé, 2004). Barcelona's aquifers suffered heavy water extractions from the XIX^{th} century until the 1970s when many industries migrated from the city to other areas. These continuous decreases in groundwater extractions, at the same time, have caused increases in water levels (Vázquez-Suñé et al., 2004). Currently, several aquifers have mainly secondary uses, such as garden irrigation and maintenance or street cleaning (Jurado et al., 2014).

The information available in the HYDOR database was obtained in different formats (paper, images and digital formats) from several hydrogeological (*e.g.*, Vázquez-Suñé et al., 2004b), hydrochemical (*e.g.*, Jurado et al., 2014), geological (*e.g.*, Velasco et al., 2012a; 2012b and Velasco, 2013) and civil engineering works, which can be used for future hydrogeological models, re-interpretations, and comparisons. Thus, the database stores over 1443 boreholes, 3700 groundwater points and 19 weather stations (for additional information on the data and their sources see Velasco 2013). In particular, in terms of pumping test data, the Barcelona database includes 70 points related to 32 pumping tests distributed in the study area (Figure 3.7).

The analysis and the interpretation of the pumping tests were performed following the next steps, which can be used as a methodology to apply the software presented in this chapter.

The first step was analysed and visualised the spatial variability of the hydraulic parameters. *MJ-Pumpit* was used to evaluate the availability of pumping tests in the selected area for the selected time period (1960 - 2014). As result, we automatically obtained necessary information to check which tests have been already interpreted (their hydraulic parameters are stored in the database) and which of the pumping tests should be interpreted or re-interpreted to ensure their quality.

The analysis and interpretation of the selected pumping test data was the later step. *MJ-Pumpit* (see section 3.2) was used to interpretate the selected pumping tests data. To facilitate the analysis, the optimal models used to obtain hydraulic parameters were selected with the aid of the automatic plots semi-log, log-log of measure data and diagnostic plots. In this case, Figure 3.2 illustrates the Barker model (Baker, 1988) with the initial values of transmissivity, storativity, well pumping storage and skin factor. After the evaluation of the results and optimal fit (Figure 3.3), the parameters values obtained were exported into the GIS environment to visualise them together with old interpretations and other tests.

Finally, the last step was to create variability map of the parameter analysed. The *Hydrogeological Test Analysis Tools* (section 3.2) had allowed creating maps of the parameters analysed. In this application was analysed the regional transmissivity (Figure 3.6), where the aquifer heterogeneities are displayed.

Furthermore, with the aid of infrastructure layers (*e.g.*, metro lines) superimposed with geology, pumping test interpretation is possible to understand the complexity of the study area at different scales, obviously, depending on the layer resolution. In particular to the MAN1 point, the test is close to an underground tunnel (forty meters approximately from MAN1 point). This kind of underground impervious structures modify the groundwater flow pattern and, therefore, the structure could be considered as a boundary condition during the test. However, compared with other tests made near the zone (MAN2 and MANS), its transmissivity values are not considerably different from each other (396 ± 74 m^2/day).

In summary, the three steps, (a) analyse and visualise the spatial variability of the hydraulic parameters, (b) analyse and interpret pumping tests with MJ-Pumpit and (c) create spatial transmissivity variability map, have been applied for the rest of the pumping tests without previous interpretations, demonstrating its usefulness, efficiency and easily understandable manner without neglecting the heterogeneity of the groundwater environment. Nevertheless, in zones where no pumping test data are available in the database, it is possible to complete the analysis with other tools including permeability tools (HEROS toolbar, Velasco et al., 2012), which can estimate the hydraulic conductivity through textural information.

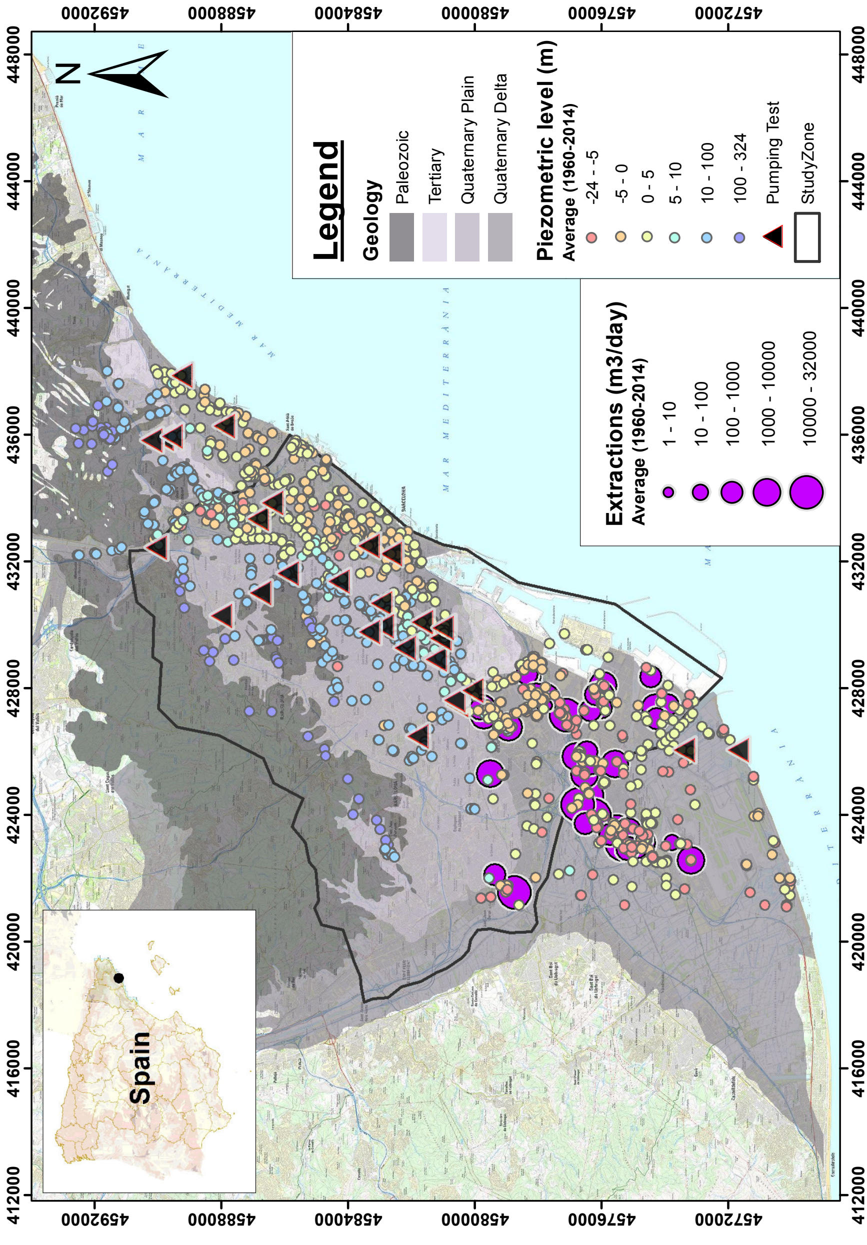


Fig. 3.7 Geological map of the Metropolitan Area of Barcelona (Spain) (IGC, Institut Geològic de Catalunya) with extractions, piezometric levels and pumping tests stored in the HYDOR database.

3.4 Conclusions

An increasing amount of data in hydrogeological studies such as pumping tests requires standardised management for an optimal analysis. These data that are collected (spatial and non-spatial dependent) can easily be integrated and handled by GIS environments. Moreover, the use of GIS platforms reduces the uncertainty of hydraulic parameter estimation because of an accurate knowledge of the aquifer geometry and hydraulic boundaries.

In the present work, we have developed a set of tools for analysing pumping tests into a GIS environment to support pumping test analysis and groundwater resource management. The software integrates different instruments to collect, manage, analyse, process and represent data derived from pumping tests analyses, through (i) an interface that improves interaction with a pumping test interpretation computer code in a platform that has extensive tools for the analysis and visualisation of data (MJ-Pumpit), (ii) tools for advanced temporal-spatial analysis built in a GIS environment (HYYH Toolbar) and (iii) a geospatial database to maintain complete control of data involved in pumping tests (HYDOR database). This set of modules constitutes an operational and user-friendly analysis platform, which exploits several analysis and visualisation tools to ensure optimal results in hydrogeological studies.

With the aid of the HYDOR database, it is possible to standardise and manage the necessary datasets to generate a complete groundwater study in an easy manner. In the same way, pumping test data management can be handled simply by using a screen selection in ArcMAP and by directly exporting the data to MJ-Pumpit. It allows the user to perform additional interpretations by considering different models, accurate geometries (*e.g.* boundary conditions) of the study area and built-in plotting tools. The interconnection among the MJ-Pumpit, GIS environment and database ensures the proper management of a hydrogeological project.

Finally, these tools are being implemented as teaching subject and applied in widely environmental projects such as groundwater management, civil works, mining, among others. Hence, these applications have been useful to fix some bugs, to improve and to maintain them in latest versions of ArcGIS. Appendix A shows more in detail registrations of these tools and institutions where have been or are being used, to our knowledge.

Chapter 4

On the reliability of hydraulic conductivity equations based on grain size data: new perspectives for scalable applications

Based on: Criollo, R.; Vázquez-Suñé, E.; Sánchez-Vila, X.; Dentz, M.; Riera, J.; Velasco, V. On the reliability of hydraulic conductivity equations based on grain size data: new perspectives for scalable applications. To be submitted in Engineering Geology (2019).

Summary

Hydraulic parameters are heterogeneous regardless of the scale of observation. The scale of interest is problem dependent, and local values of hydraulic conductivity (K) are scale dependent. Different methods exist to obtain K at the local scale. These methods can be classified into two main groups based on whether K is directly obtained from the interpretation of laboratory permeameter tests of undisturbed samples or indirectly estimated from empirical correlations based on the grain size distribution (with sample disturbance prior to testing). Many empirical formulations have been devised over the last 120 years to derive estimates of K from particle size distribution information, whether based on simple representative sizes or on the full distribution of grain sizes. We present a different classification of methods by sorting the methods into six groups based on the type and amount of data required for each empirical formula. The estimates of K using 51 different empirical methods (involving empirical equations, pedotransfer functions and a neural network) were compared with the K values from 28 permeameter tests involving samples composed of loose sand with a very low clay content. The results were used to rank the methods and select the optimal empirical equation as a function of the amount and type of information available. Additionally, an upscaling analysis of these equations was performed to ensure the applicability of each formula at larger scale.

4.1 Introduction

Hydraulic conductivity, K, is arguably the most significant parameter in hydrogeology. All groundwater studies require proper site characterisation in terms of K values at the local scale. Local values are representative of a given volume. Changes in this volume inevitably lead to variations in the corresponding representative values; therefore, K values are scale dependent.

At the decimetre scale (*i.e.*, corresponding to ten to a few tens of centimetres), the K value at a given point depends on the characteristics of the medium, such as the degree of cementation, porosity, grain size distribution, grain shapes, pore sizes and interconnections, as well as the weathering and compaction processes the soil has experienced. This size is arguably the smallest that can be meaningfully used to define K because a scale smaller than this can, in some cases, violate the concept of the representative elementary volume (REV), *e.g.*, Bear (1972). Decimetre-scale values can be used as the first step in any upscaling process leading to representative K values at larger scales (*i.e.*, scales that could be used in subregional numerical models) (*e.g.*, Gleeson et al., 2016b). Depending on the stratigraphic structure, conductivity can be locally considered a scalar or a symmetrical second-order tensor (termed $\boldsymbol{k}$).

A number of studies have investigated the relationships between the values of K at different support volumes (see, *e.g.*, the reviews by Chapuis, 2012; Bouwer and Rice, 1976; Renard and De Marsily, 1997; and Sánchez-Vila et al., 2006). A more general approach is to consider the K value for a given support volume a spatial random function (SRF) with a predefined multivariate probability density function (PDF) (*e.g.*, Dagan, 1989 or Gelhar, 1992). Assuming a simplified model, the SRF can be defined by a univariate PDF and a correlation function. The upscaling process results in a change in the multivariate distribution, so that the arithmetic mean and variance of the point K values decrease and the integral distance increases. In general, scalar values at the local scale will produce tensorial upscaled values of K (*e.g.*, de Dreuzy et al., 2010; Gelhar and Axness, 1983).

Any hydrogeological model should be based on a thorough description of the spatial distribution of conductivity values at the scale of interest (*e.g.*, the scale of discretisation in a numerical model). However, the amount of available hydraulic conductivity data at any site is generally small with respect, for example, to the number of cells in a numerical model. Moreover, existing data may represent different support sizes, which can even differ from the most suitable support size for a given model. Thus, there is a need to either extrapolate local values from a few measurements with geostatistical methods or rapidly obtain values of K in an inexpensive manner.

Hydraulic conductivity values at the decimetre scale can be obtained by different approaches that can be classified as direct or indirect methods.

Direct methods

Direct methods include different field and laboratory approaches to estimating K. Field methods usually consist of a suite of hydraulic tests, *i.e.*, long- or short-term pumping tests, slug tests, flowmeter measurements, air permeameter measurements, etc. In these tests, K is derived from the recorded hydraulic response at one or more observation points. Each approach relies on a particular interpretation method (usually based on a number of simplifying approaches) and provides a value that is representative of a specific support volume (*e.g.*, Sánchez-Vila et al., 2006).

Laboratory tests include different approaches that mimic Darcy's experiment using either constant-head or falling-head permeameters, where K is obtained from the direct application of Darcy's law. The representative scale of the value obtained is that of the size of the sample.

Indirect methods

Indirect methods derive K values from indirect measurements. One subgroup involves the use of different geophysical techniques. However, in this work, we are interested in K estimates derived from the grain size distributions of individual samples. The earliest studies on this topic can be traced back to Hazen (1892), who provided the first empirical relationship between K and an equivalent (or representative) grain diameter. Since then, many authors have proposed variations to this relationship, such as by including the sample porosity (Slichter, 1899, in Custodio and Llamas, 1984). Subsequently, different empirical formulations have been developed incorporating porosity (Terzaghi, 1943) and soil characteristics (Marshall, 1958 in Cronican and Gribb, 2004).

In addition to these methods, estimates of K from data recorded in soil surveys, such as the percentages of clay, silt, and organic matter; the particle size distribution; and the water retention, can be obtained using pedotransfer functions (PTFs). Although the initial work on this topic can be traced back a century (Briggs and Shantz, 1912), it was formalised many years later (Bouma, 1983; Hamblin, 1991). In short, K values are obtained from a large number of parameters based on regression equations. These equations have evolved from the early linear models to the more recent nonlinear models. In the past few years, principal component analysis has been performed as a common alternative to obtaining such regressions (for further information, please see Wösten, 1997 and Wösten et al., 2001). An alternative to regression equations is using optimisation learning models, such as artificial neural networks (ANN), which have been increasingly reported in the literature (*e.g.*, Schaap et al., 1998; 2001, Rogiers et al., 2012).

The main advantage of indirect methods is their capability of providing simple, semi-automated and inexpensive values of hydraulic conductivity at the scale of the corresponding soil sample (Vukovic and Soro, 1992; Kasenow, 2002, Merdun, 2010; Vienken and Dietriech, 2011). This simplicity may be why so many empirical formulas have proposed over the past hundred years. A review of a subset of formulations is included in the Appendix D.

In some cases, using the same data on the grain size distribution of a given sample, individual empirical formulas can provide different estimates of the local K. Therefore, a number of authors have contributed to evaluating and comparing such formulas. For instance, Vukovic and Soro (1992) compared estimates of K obtained using ten selected empirical formulas with those derived from pumping test analyses with the intention of providing a range of applicability for such formulas. However, the effects of the different support scales between the granulometry and hydraulic tests were not considered. In a similar study, Cheong et al. (2008) compared the K values obtained based on aquifer tests (pumping and slug tests) with numerical model results and the result of five empirical formulas based on grain size analysis. Carrier (2003) compared and evaluated two empirical formulas (Kozeny-Carman and Hazen) and concluded that the Hazen formula is less accurate than the Kozeny-Carman formula. Vienken and Dietriech (2011) compared the estimates from seven frequently used formulas with K estimates from high-resolution direct push tests (DPSTs). Similarly, Matthes et al. (2012) evaluated six of the most commonly used empirical equations considering the measurement accuracy and the determination of K from grain size data from seven samples. The results showed that the variability in K was highly influenced by strong differences in K estimates among formulas. Chapuis (2012) evaluated forty-five predictive methods for the saturated hydraulic conductivity of soils. Most predictive methods were calibrated based on a permeability test and displayed the typically error for these types of tests. As a result, a range of applicability was established for different types of soils. Wagner et al. (2001) compared eight different PTFs and found that the best correlation was provided by the Wösten method (on average) for the prediction of the unsaturated hydraulic conductivity.

Another method used to estimate values of K is by neural network models, which can be used to develop PTFs of soil hydraulic characteristics. Cronican and Gribb (2004) compared their equation, the formula of Rawls and Brakensiek (1989) (obtained from multiple linear regression) and a neural network based-model (Rosetta, Schaap et al., 2001). They concluded that the Rawls and Brakensiek and ANN model estimates were one to two orders of magnitude greater than the measured K values, and their equation is only applicable for soils with sand and clay percentages within a given range. Merdun et al. (2010) evaluated the K estimates from PTFs obtained by cascade forward network (CFN), multiple linear regression (MLR) and seemingly unrelated regression (SUR) models. They found that the CFN displayed the highest accuracy in the estimation of K. The ANN and multiple regression analysis (MRA) models developed by Erzin et al. (2009) yielded accurate hydraulic conductivity values for fine-grained soils. Rogiers et al. (2012) evaluated two data-driven modelling methods (MLR and ANN methods) that use grain size distribution data to predict K values. Their outcomes displayed reasonable accuracy for predicting the aforementioned parameters with MLR combined with principal component analysis. Nevertheless, the GLUE-ANN (Generalised Likelihood Uncertainty Estimation ANN) proved to perform slightly better than the MLR method. Thus, the uncertainty associated with GLUE-ANN estimates was considerably less for several stratigraphic units than was that based on the MLR method.

The objective of the present work is twofold. First, to evaluate the applicability of common empirical formulas, PTFs and ANNs based on grain size analysis are calibrated by comparison with 28 laboratory permeameter tests conducted with sandy samples to obtain a ranking of the different methods. The compilation of methods can be seen as an extension of the works of Kasenow (2002), Cronican and Gribb (2004) and Chapuis (2012). However, the inclusion of the results of the permeameter tests permitted the ranking of the different methods as a function of the amount of information included in each individual formula. Second, statistical analyses were performed to upscale K and identify the methods that can be confidently applied at large field scales in real case studies.

In section 2, information about the sampling procedure, permeameter tests and grain sieving analysis is reported. Section 3 presents the different empirical methods, which are classified as a function of the information required by each formula. Additionally, the statistical approach used to compare the methods and to analyse their upscaling behaviour are presented. A comparison of the performance of the formulas and their upscaling behaviour are given and discussed in section 4. Finally, a set of conclusions and further analyses are described.

4.2 Materials and Methods

Study site

The field site is located in the Llobregat Delta, close to the city of Barcelona (Spain) (Figure 4.1A). The delta body is composed of four non-formal units (from bottom to top): (1) blue marls and clays containing marine Pliocene fauna; (2) Pleistocene muddy conglomerates overlain by a coarse sequence of gravels interfingered basinwards with sands; this unit constitutes the lower, semiconfined aquifer; (3)

Holocene marine silts interfingered with proximal gravels; and 4) the most recent unit, mainly formed by clean, well-sorted sands; this unit forms the upper, unconfined aquifer. Thorough geological and hydrogeological descriptions of the area were provided by Cabello et al. (2007) and Gámez et al. (2009).

The study site consists of a shallow excavation 140 m long, 40 m wide and 8 m deep (Figure 4.1B). Five stratigraphic logs were collected and subsequently analysed. The log records indicated clean sands with small gravel contents and shallow lacustrine-marsh mudstone and peat (Cabello et al. 2007).

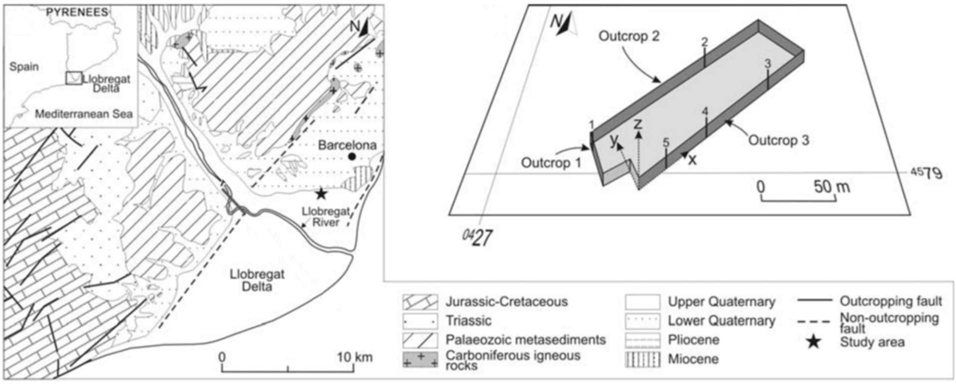


Fig. 4.1 Location and geological map of the Llobregat Delta and its surroundings (left). Map of the study area with the different outcrops (right) (Modified from Cabello et al., 2007). Numbers 1 to 5 correspond to the locations of logs where undisturbed samples were extracted. Coordinates are given in kilometres in the Universal Transverse Mercator (UTM) projection.

Undisturbed sampling

Standard penetration tests (SPTs) with depth correction were performed to collect undisturbed samples. Although the soil was mainly composed of sand, 20-30 evaluations suggested that the soil was sufficiently compact for obtaining unaltered samples (see Alonso et al., 2000).

Undisturbed sampling involved first constructing PVC cylinders with a 50-mm inner diameter, 5-mm thickness, and sharp edges. The cylinders were used to obtain 10-cm long samples with minimal disturbance. The cylinders were driven into the soil manually, after which each sample was carefully removed and wrapped in plastic film. All samples were collected parallel to the layered structure. The samples were preserved in a wet chamber until the permeameter tests were performed. Overall, 28 unaltered samples were used in this work after rejecting samples with evidence of alteration.

Laboratory testing

All samples were first subjected to falling-head permeability tests. This approach is a common laboratory test method used to determine the permeability of relatively fine-grained soil (for further information, see Head, 1982). Each individual sample was placed in the testing device and sealed with silicon grease to ensure tightness. Then, the sample was saturated with distilled and degassed water. The head variations were recorded during the test and plotted versus time in a semilog plot. Each test was repeated 5 times, and the average slope was computed. The hydraulic conductivity was then derived from the average slope.

To obtain the sample bulk porosity, each sample was resaturated and weighed. Next, samples were dried in an oven for two days at 90 ºC. Finally, the dried samples were weighed to calculate the volume of water under saturated conditions. From the known total volume of the sample, the bulk porosity was inferred as the volumetric ratio. Permeability and porosity tests were performed at the Geotechnical Laboratory of the Department of Geotechnical Engineering and Geosciences (UPC-Barcelona Tech).

The grain size distribution is typically obtained by plotting the percentage of solid mass (p) smaller than size d (mm), as determined by sieving and hydrometer tests, versus the decimal logarithm of d ($\log(d)$) (Chapuis, 2012). However, it is possible to obtain more precise values using the frequencies of the grain sizes based on the Coulter counter method (Walker et al., 1974; Konert and Vandenberghe, 1997; Eshel et al., 2004; Blott and Pye, 2006), which yields numerical values of the grain diameter (d_{10}, d_{50}, d_{60}, etc.) and the percentage of each material (sand, silt and clay).

The grain size analysis was performed with a *Coulter LS 100* system located at the Sedimentology Laboratory of the Geology Faculty of Universitat de Barcelona. Each sample was well mixed and homogenised to obtain a representative subsample of each individual sample. A Coulter counter measures the angular distribution of diffracted light at different time intervals, yielding the full particle size distribution curve for each sample. The measurement time in each sampling interval was 60 seconds without interruption. These results were averaged to obtain the final size distribution for the given sample. The Coulter counter system provides grain size readings ranging between 0.38 and 1000 μm. Fractions of grain sizes above 1 mm were determined by sieving the full sample and weighing the particles collected.

Pumping tests

Five pumping tests in the study area (two for shown general values of the aquifer and three executed near the excavation site) were performed with constant flow based on FCIHS guidelines (www.fcihs.org). The wells and piezometers used during these aquifer tests were completed in the upper aquifer (18 metres), and their diameters were 0.1 to 0.2 metres. The maximum static water depth was 7 metres before starting each pumping test. After dewatering to obtain steady-state water levels, the pump was stopped to measure the recovered water values until static initial values were observed. Pumping tests were interpreted with Ephebo (UPC, 2002) and MJ-Pumpit (Criollo et al., 2016). Pumping tests characteristics and their interpretation results are summarised in Table 4.1.

Test	Point name	Diameter (m)	Distance from well to centre of the study zone (m)	Flow (m3/day)	Static water depth (m)	Hydraulic Conductivity (m/day)
Desv. Río	Desvrio	0.2	5109	2000	3	12.4
Pryca	Llob1	0.1	3458	2000	3	18
Well4	Piezo4	0.2	20	260.4	7.67	30
Well5	Piezo5	0.15	20	380.2	7.36	38.6
Well6	Piezo6	0.15	20	475.2	7.44	48

Table 4.1 Characteristics of the pumping test used in the comparison with upscaling results.

4.3 Empirical methods: classification and performance assessment

Classification of empirical methods

The study begins with the compilation of 51 empirical equations to estimate K using grain size data with or without additional information. These methods were initially filtered to discard duplicates, errors and equations with no widely available information to be applied. A total of 41 methods were finally considered. We classified these methods into six families (i to vi) depending on the type and amount of information involved in the estimation. All the equations are included in the Appendix D. The current classification of the methods into families is shown in Table 4.2. The final families selected are as follows.

	Porosity is not directly included in the estimation formulae	**Porosity is included explicitly in the estimation formulae**
Empirical estimation from sieve data	**Family i** Hazen, Harleman, Shepherd, Hazen-Tylor, Kenney, USBR, Uma, Zunker*, Hullvert-Feben*, Sullivan-Hertel*	**Family iv** Breyer, Hazen-Tison, Hazen-Schoeller, Slichter, Terzaghi, Kozeny, Kozeny-Carman, Schönwalder, Rumer, Bakhmeteff, Bakhmeteff-Feodoroff, Sperry-Pierce, Chapuis (a), Capuis (b), Krüger, Choo*, Irmay*, Kozeny-Fair-Hatch*
Pedotranfer functions	**Family ii** Cosby, Campbell, Jabro, Li, Dane-Puckett, Puckett, Saxton, Rosetta	**Family v** Vereecken, Cronican-Gribb, Brakensiek, Rawls-Brakensiek, Mbonimpa
From statistical correlations	**Family iii** Krumbein-Monk, Alyamani-Sen, Bloemen*	**Family vi** Fair-Hatch, NAVFAC, Shahabi, Ahuja*, Marshall*, McDonald*, Barr*

Table 4.2 The classification of the methods used in this study. See the Supplementary Material for the actual formulas adopted in the study. Methods marked with (*) have not been evaluated.

Family i comprises a suite of empirical equations that use only equivalent diameter values obtained from grain size analysis. These equations involve the least amount of information. K is mainly governed by a diameter that is representative of small grains. Most equations in this family also incorporate an additional equivalent grain size representative of middle-range grain sizes.

Family ii includes linear or nonlinear functions involving all or some soil texture, silt or clay percentage, bulk density, organic matter content, and saturation information at different entry pressures. Most of these methods can be included in the general family of PTFs, with equations involving low- or high-order polynomials of different parameters. In this family, we have also included a method based on an ANN implemented with Rosetta (Schaap et al. 2001). Rosetta combines a neural network with the bootstrap method to calculate both K values and uncertainty estimates. We used two different networks depending on whether bulk density (BD) information was included. We termed these methods Rosetta-BD and Rosetta-no BD.

Family iii includes methods based on statistical correlations (standard deviations) with grain sieving data and some experimental constants. One of these methods assumed that the ratio of d_{10} to d_{50} is proportional to K, and it preserves the relation between d_{10} and K proposed by Hazen (1892) (Alyamani and Sen, 1993). Such constants are linked to qualitative descriptions of grains, including the roughness or shape, and are shown in the Appendix D.

Families iv, v, and vi include equations similar to those in families i, ii and iii, respectively, with the unique difference that they explicitly include the value of porosity in the estimation of K.

Some constraints for the equations presented in the Appendix D must be considered. In the literature, some equations were devised to obtain the hydraulic conductivity K [L/T], whereas others provide estimates of intrinsic permeability k [L^2]. Accordingly, it is therefore necessary to proceed with caution and avoid any confusion between the two terms. Its consistency is favourable when temperature effects on the fluid viscosity (Vienken and Dietrich, 2011). Both parameters are related to the density and kinematic viscosity of water. To avoid confusion, the Appendix D shows all equations converted to K using the specific weight (γ) and kinematic viscosity (μ) of freshwater at a temperature of 20 ºC, such that $\gamma/\mu = 9.76 \cdot 10^{-6}$ $(\mathrm{m} \cdot \mathrm{s})^{-1}$ (Schneebeli, 1966). The use of different units is another source of discrepancies because many equations were proposed before an international system of units was established as the standard. To simplify and facilitate the relevant tasks and improve the clarity of this manuscript, all the compiled equations were converted to parameters based on metres and days. Additionally, some formulas use qualitative descriptions. To avoid ambiguity, we used the USDA classification, where "sand" describes the fraction of the soil with grains of an equivalent diameter exceeding 50 μm, "silt" is that with an equivalent diameter between 2 and 50 μm, and "clay" is that with an equivalent grain diameter lower than 2 μm. The samples used in this study were mainly composed of sands, with a small portion of silts and negligible clay contents. Thus, the results obtained in this work should not be extrapolated to silty or clayey soils. Additionally, the different sample characteristics must be considered.

The last point of caution is that some of these methods have been used for decades, and many have since been proven inaccurate in the transcription of the equations from the original manuscripts. In most cases, these issues can be attributed to the improper correction of units. We performed a detailed analysis of the original sources to recover the original formulas and provide updated empirical parameters that allow for a direct comparison of K estimates from different methods. A list of the inaccurate transcriptions we found in the literature has been compiled in the Appendix C.

Statistical analysis

The assignment of equations into families aids in (i) classifying the different families and the different equations within each family to determine which methods provide K estimates that best fit the actual measured hydraulic conductivity values from the permeameter tests and in (ii) determining which methods can be confidently applied at large field scales.

The permeameter test results and the most relevant parameters characterising the grain size analysis are presented in Table 4.3. For instance, some samples exhibited 0% clay fractions, potentially due to the detection limit of the Coulter counter. We adopted a minimum clay content of 0.1% in each individual sample to use the formulas that involve the logarithm of the clay fraction. This selection has some implications that apply to specific methods in families ii and v.

Moreover, the parameters used in the empirical formulas (Families ii, iii, v, and vi), excluding those with statistical correlations (Family ii), include the porosity (Φ); the percentages of sand (S), silt (u) and clay (c); the bulk density (B_d); and the representative diameter sizes d_{10} and d_{60}. The values corresponding to each sample are listed in Table 4.3.

Statistical description of K values from the permeameter tests

Because the first goal of this study is concerned with the estimation of individual K values, we analyse the univariate K statistics obtained from the permeameter tests. The basic statistical soil characteristics obtained from the laboratory tests are given in Table 4.3. Figure 4.2 presents a box plot of K values and a relative frequency plot of log conductivities. Most individual values range from 0.01 to 10 m/day, and the

Sample	*K* *(m/day)*	*Φ* *(-)*	**S** **(%)**	**u** **(%)**	**c** **(%)**	**Bd** **(g/cm3)**	**d10** **(μm)**	**d60** **(μm)**
1	*3.6*	*0.54*	95.70	4.30	0.1*	1.2	94	180
2	*17.1*	*0.25*	97.80	2.30	0.1*	1.94	146.8	280
3	*23.6*	*0.33*	81.90	0.1*	0.1*	1.75	265.1	1800
4	*28.6*	*0.3*	95.30	1E-5 *	0.1*	1.83	193.8	1400
5	*17.6*	*0.35*	79.10	18.60	2.1	1.68	19.38	300
6	*12*	*0.4*	95.00	4.10	0.1*	1.56	90.63	200
7	*9.2*	*0.42*	98.30	2.30	0.1*	1.52	173.5	350
8	*14.7*	*0.41*	97.20	3.60	0.1*	1.54	144.7	300
9	*20.3*	*0.36*	86.40	0.1*	0.1*	1.66	212	1500
10	*16.1*	*0.37*	93.80	6.50	0.1*	1.64	138	1600
11	*14.9*	*0.33*	82.40	9.50	0.3	1.74	65.89	1200
12	*2.9*	*0.42*	85.60	13.30	1.3	1.5	42.62	160
13	*0.7*	*0.41*	93.80	26.30	2.5	1.55	23.21	120
14	*1.7*	*0.4*	80.20	18.9	1.8	1.55	44.46	110
15	*7.5*	*0.46*	93.10	6.1	0.1*	1.41	79.82	200
16	*20.7*	*0.33*	99.70	0.1*	0.1*	1.73	220.6	1500
17	*29.8*	*0.3*	100.00	0.1*	0.1*	1.83	364.1	2000
18	*15.1*	*0.32*	97.90	4.0	0.1*	1.77	134.8	280
19	*31.8*	*0.42*	99.70	0.7	0.1*	1.51	239.4	500
20	*0.9*	*0.33*	83.50	14.8	1.8	1.73	24.51	280
21	*6.2*	*0.46*	93.70	6.0	0.1*	1.41	79.92	170
22	*4.9*	*0.43*	94.40	5.8	0.3	1.49	85.75	250
23	*2.7*	*0.47*	80.50	18.8	1.4	1.38	41.82	120
24	*2.6*	*0.42*	84.30	15.1	1.0	1.5	43.07	160
25	*8.2*	*0.42*	95.50	4.3	0.1*	1.52	102	210
26	*10*	*0.39*	96.70	2.1	0.1*	1.59	129.3	250
27	*11.9*	*0.34*	100.30	0.1*	0.1*	1.71	169.2	600
28	*6.9*	*0.41*	96.20	3.9	0.1*	1.54	129.7	260
mean	*12.54*	*0.39*	91.40	9.25	1.4	1.60	122.80	606.80
St. deviation	*9.31*	*0.06*	6.95	7.03	0.73	0.17	86.02	620.40

Table 4.3 Representative parameters of the samples considered in this study. The values approximated to avoid a zero value are marked with an asterisk (*). K is the hydraulic conductivity in [m/day] estimated from the permeameter tests; Φ is the measured porosity; s, u, and c are the sand, silt and clay percentages, respectively; Bd is the bulk density in [g/cm3]; d_{10} is the particle size for which 10% of the soil is finer in [μm]; and d_{60} is the same as d_{10} but for 60%.

highest *K* value is 31.8 m/day. The distribution has a significant positive skew, with a log-*K* variance of 81.6.

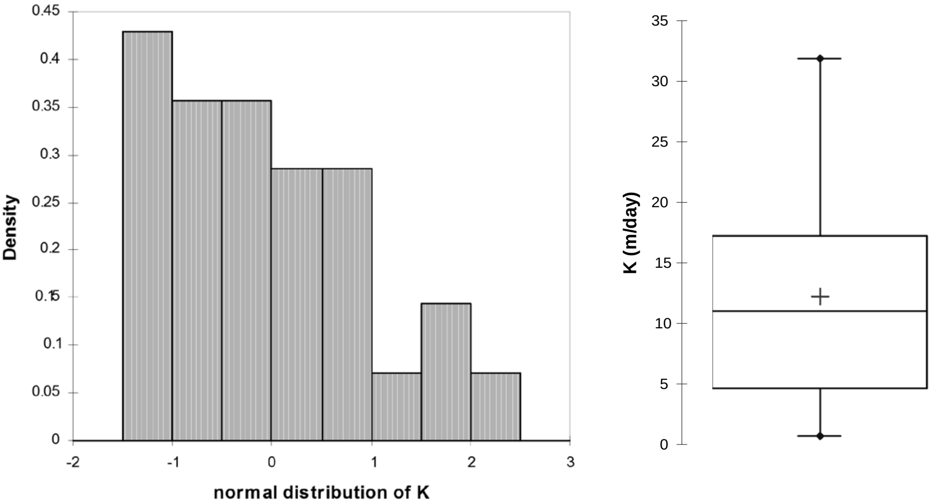


Fig. 4.2 Relative frequency histogram of log-K values obtained from permeameter tests of the 28 samples (left). Corresponding box plot of the K values showing positive skewness (right).

Selecting the statistical indicators of the performance of each empirical formula

The effectiveness of the different empirical formulas in reproducing the *K* values obtained from the permeameter tests was quantified in statistical terms. Fifty empirical methods and one ANN used to calculate conductivity and intrinsic permeability were found in the literature (see Appendix D). After prescreening, only 41 of these methods were used because the remaining methods include parameters that are too specific and not widely available, which makes them of little use in most real field applications. Due to the large range of *K* values, it is easier to work with log-K values. To quantify the goodness of fit of a given method, we sequentially studied each formula by comparing the estimated *K* values with those obtained from the permeameter tests in terms of the logarithmic error (E_{log}) and Pearson's correlation coefficient (R), which are defined as follows:

$$E_{log,j} = \frac{1}{n} \cdot \sum_{i=1}^{n} log(K_{p,i}) - log(K_{g,i,j}) \tag{4.1}$$

$$R_j = \frac{Cov(K_p, K_{g,j})}{\sigma_{K_p} \cdot \sigma_{K_{g,j}}} \tag{4.2}$$

where n is the total number of samples; $K_{p,i}$ is the hydraulic conductivity of sample i measured in the permeameter test; $K_{g,i,j}$ is the conductivity of the i^{th} sample calculated with the j^{th} empirical formula; and σ and Cov are the standard deviation and covariance of the corresponding arguments, respectively. We consider an empirical formula j to produce acceptable results when $E_{log,j}$ is less than $|0.5|$ (indicating that the mean error is below half an order of magnitude) and R_j is larger than 0.6. These values are selected after checking the results of all methods and are arbitrarily used to select the methods that best reproduce better the experimentally measured values (*i.e.*, other values could have been chosen).

Upscaling *K* for large-scale applications

Because the second objective is to determine which methods in each family can be confidently applied at a larger field scale, further statistical analyses are performed to upscale K and validate the application of each formula in a real case study. To validate this approach, the upscaling results are compared with pumping tests results obtained close to the study area. To represent the study area, and a novelty in the upscaling methodology, an integrated sample (named CAR-INT) is created from the Coulter counter data from each sample. All the equations were applied based on this new integrated sample.

The theoretical equivalent K were obtained with the simplified power average method proposed by Journel et al. (1986):

$$K_a = K_p^{\omega} \tag{4.3}$$

where K_p is the hydraulic conductivity and ω is an exponent between -1 and 1. The minimum value (ω = -1) corresponds to flow perpendicular to a layered structure (based on the harmonic average), and the maximum value ($\omega = 1$) corresponds to flow parallel to a layered structure (based on the arithmetic average). Note that when $\lim_{\omega\to\infty}(K_p^{\omega}) = 0$, K_a corresponds to the geometric average.

Applying last equation in the k measured in the laboratory of each sample we obtain limits of the upscaling and the theoretical equivalent K (the geometric mean) of our study zone. Afterwards, results of each K equation using the integrated sample (CAR-INT) have been compared with limits of apply equation (4.3) and with pumping tests.

The discussion and results of the statistical analysis are given below (Section 4), and the conclusions are presented in Section 5.

4.4 Results and discussion

We recall that our objectives are twofold: to evaluate the goodness of fit of the K equations by applying indicators (4.1) and (4.2) and to evaluate how the equations that satisfied the first objective can be applied in real cases. As stated in the introduction section, several works have investigated the direct relation among K equations based on grain sizes determined by aquifer tests but have not considered scale effects in this relationship. Hence, our second objective has been analysed the effects of the upscaling in the methods compiled. To achieve that, we unify all the samples in order to create an integrated and unique sample. Then, this integrated sample has been applied in the K equations compiled. Afterwards, the equation (4.3) is used to upscale the values of the study area, in order to compare both results. To validate the upscaling, we compare them with the K values obtained from the pumping test.

Comparison of the statistical indicators of the performance of each *K* formula compiled

The hydraulic conductivities of different samples were estimated with the empirical methods and compared to the respective values obtained from the permeameter tests using indicators (4.1) and (4.2). Sample 5 was excluded from the statistical analysis due to possible data errors leading to no formula obtaining a satisfactory estimate of the corresponding conductivity. Notably, the estimates were not of the same order of magnitude as the measured values. For completeness, sample 5 is included in all figures, which show the conductivity values obtained for each sample based on different empirical methods.

Table 4.4 contains the logarithmic error (4.1) and Pearson's R (4.2) results for the different empirical methods and the six families defined in Section 3. The values fulfilling the goodness of fit criteria established in the previous section ($E_{log,j} < |0.5|$ and $R_j > 0.6$) are given in bold.

Family i: The methods proposed by Hazen (Custodio and Llamas, 1984) and Harlemann (Schwartz and Zhang, 2003) provide the best performance, with |E| < 0.20 and R > 0.81. As noted in the (Appendix C), a number of values have been proposed for the empirical coefficient of the Hazen method, also called the empirical Hazen coefficient (Carrier, 2003). A value of 100 (Custodio and Llamas, 1984) and other values have been used (see Kenney et al. (1984) or Uma et al. (1989) for some examples). All the methods are detailed in the Supplementary Material (Appendices C and D).

For the Shepherd method (Shepherd, 1989), we performed a sensitivity analysis with respect to the empirical parameters (see Table 4.5 for all the sensitivity analyses). As Figure 4.3A shows, four of these sets of parameters correspond to values proposed by the author (Shepherd, 1989): one for dunes, one for beach sand and two extreme values. The absolute logarithmic error between the different parameter sets varies over two orders of magnitude, with a coefficient of variation of -1.11 (Table 4.6). This finding demonstrates the importance of the shape of the material in this method and suggests that the Shepherd method can be unstable if sufficient heterogeneities are present in the studied materials.

Family ii: The values obtained based on statistical comparisons (4.1) and (4.2) are not as good as those produced for the methods in Family *i.* Most of the methods provide mean errors and correlation coefficients outside the range established ($E_{log,j} <$ $|0.5|$ and $R_j > 0.6$); for example, the Rosetta method (see Figure 4.4) displays a mean error of 2.48. The performance of all the methods is described in Table 4.4.

	Method	**Elog**	**R**
Family i	**Hazen (c=100)***	**0.006**	**0.81**
	Harleman	**0.20**	**0.81**
	Shepherd	**-0.25**	**0.85**
	Hazen-Taylor	**-0.06**	**0.81**
	Kenney	**0.31**	**0.81**
	USBR	**0.39**	**0.67**
	Uma	**0.24**	**0.81**
Family ii	Cosby	0.56	**0.70**
	Campbell	1.04	-0.27
	Jabro	**0.25**	-0.25
	Li	-1.48	-0.39
	Daneand Puckett	**0.07**	0.46
	Puckett	**0.36**	0.46
	Saxton	-0.61	-0.57
	Rosetta-BD	2.49	0.51
	Rosetta-no BD	2.48	**0.71**
Family iii	Krumbein-Monk	0.87	**0.69**
	Alyamani-Sen	**0.33**	0.42
Family iv	**Breyer**	**-0.08**	**0.81**
	Hazen-Schoeller	**0.43**	0.41
	Hazen-Tison	**0.27**	0.41
	Slichter	**0.33**	**0.73**
	Terzaghi	**0.21**	**0.71**
	Kozeny	**-0.29**	**0.62**
	Kozeny-Carman	**0.11**	**0.62**
	Schönwalder	-0.71	**0.73**
	Rumer	**0.17**	**0.78**
	Bakhmeteff	**-0.28**	**0.82**
	Bakhmeteff-Feodoroff	**0.08**	**0.82**
	Sauerbrei	**0.21**	**0.62**
	Krüger	**0.24**	**0.80**
	Sperry-Pierce	-1.23	**0.69**
	Chapuis (2004a)	**0.45**	**0.62**
	Chapuis (2004b)	**-0.17**	**0.63**
Family v	Vereecken	**-0.17**	-0.28
	Cronican-Gribb	0.62	**0.63**
	Brakensiek	**0.43**	-0.22
	Rawls-Brakensiek	**0.43**	-0.22
	Mbonimpa	1.07	0.24
Family vi	**Fair-Hatch**	**0.18**	**0.71**
	NAVFAC	**0.11**	0.24
	Shahabi	**0.10**	0.40

Table 4.4 Logarithmic average error (E_{log}) and correlation coefficient (R) obtained by comparing the results of the permeability methods and laboratory data. The results within the range are given ($E_{log,j} < |0.5|$ and $R_j > 0.6$) in bold.

		Values described by theauthor	**Values proposed in this work**		**Values described by the author**	**Values proposed in this work**	
Method	**Parameter name**	***Mature and rounding***	***Mature and subrounding***	***Immature and subangular***	***Immature and angulous***	***Dune***	***Beach***
Rumer	α	207	188.66	170.32	152		
Kozeny	C by Schwarz and Zhang (2003)	0.5	0.52	0.54	0.562		
	C by Schoeller (1962)	0.025	0.065	0.105	0.145		
Fair-Hatch	α	6	6.567	7.134	7.7		
Terzaghi	α	800	686.66	573.32	460		
Shepherd	a	208808	139544	70279	1014	12000	3500
	b	2.05	1.21	0.37	1.11	1.75	1.65
Kozeny and Fair-Hatch	A	1/50	0.00633	0.00599	1/175		

Table 4.5 Values for the maturity and angularity extremes proposed in a previous study and the moderate conditions proposed in this study. The effects of these values on the results are further analysed.

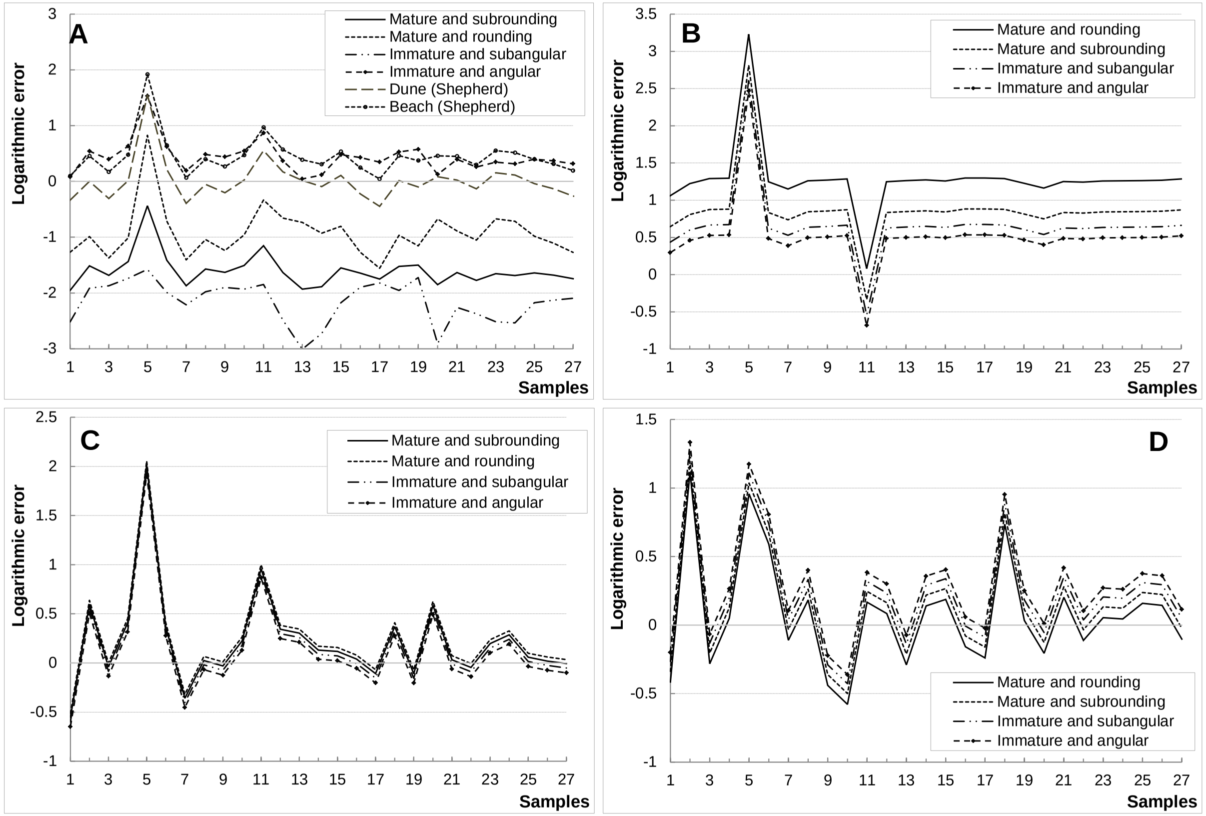


Fig. 4.3 Sensitivity study of the Shepherd (A); Kozeny-Schoeller (B); Rumer (C) and Fair-Hatch (D) methods. The values used to obtain the K results are given in Table 3.

Family iii: None of the methods included in this family provide good K estimates. The Krumbein and Monk (1943) methods meets the requirement of $R > |0.6|$, but the corresponding dispersion is high compared with other relevant methods included in other families, as is graphically shown in Figure 4.5.

Family iv: All the methods in this family except the Hazen-Tison (Schoeller, 1962); Schönwalder (1928) and Sperry and Peirce (1995) methods meet the acceptance criteria. Some methods use additional empirical parameters that are related to the maturity and roundness of grains. The authors of these empirical methods frequently propose considering two extreme cases, the mature and round case and the immature and angular case, to obtain a range of possible K values for each individual sample. In this study, we added two intermediate categories, mature and subrounded and immature and subangular, to assess the sensitivity of the estimated K values to these empirical parameters. To assess the behaviours of the methods with various maturities and angular properties, a sensitivity study was conducted for the Kozeny (1953), Terzaghi (1925) and Rumer (1969) methods using four parameter sets corresponding to the extreme values proposed previously and the intermediate values proposed in this study. The latter values are described in Table 4.5. The slight oscillation in the results of these analyses around a logarithmic error of zero indicates the importance of these properties in each method, as shown in Figure 4.4.

Family v: All the methods included in this family provide estimates that do not comply with the acceptance criteria. The Cronican and Gribb (2004) method (E_{log} = 0.62; R = 0.63) is the closest to meeting the acceptance criteria established (see Figure 4.4 and Figure 4.5 for E_{log} and R, respectively).

Family vi: Of all the methods included in this family, only the Fair-Hatch method (1933) complies with the acceptance criteria. The equation is based on dimensional considerations and experimental verification. The parameters involved in this method reflect the shape of the material, the flow (viscosity and density of water) and textural characteristics. Figure 4.3D shows how the material shape properties are used to weight the K values in the Fair-Hatch method (1933). The logarithmic error of this method is low (0.18) compared with other those of other methods in which the shape of the material plays a significant role (see Table 4.6).

In summary, the methods that meet the acceptance criteria (family level) are as follows: **Family i)** the Hazen, Harlemann, Shepherd, Kenney, USBR, and Uma methods; **Family iv) the** Breyer, Slichter, Hazen-Taylor, Terzaghi, Kozeny, Kozeny-

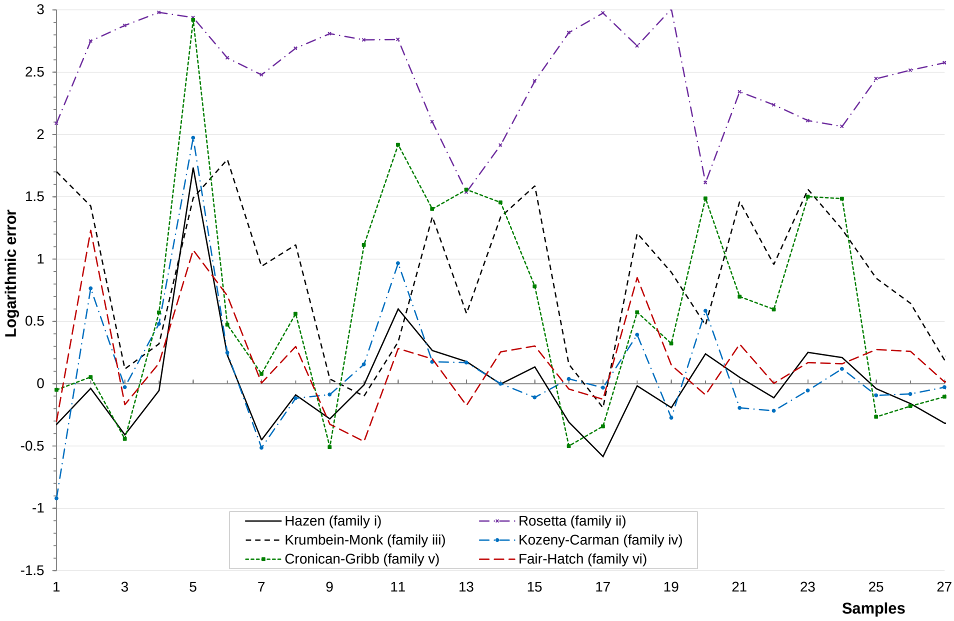


Fig. 4.4 Logarithmic errors of the major methods included in each family.

Methods	Logarithmic error mean ($E_{log\ mean}$)	Standard deviation (σ)	Coefficient of variation (V)
Rumer	0.176	0.058	**0.327**
Sheperd	-1.071	1.117	**-1.044**
Kozeny (Schwarz and Zhang, 2003)	-0.317	0.022	**-0.069**
Kozeny (Schoeller, 1962)	0.826	0.333	**0.403**
Fair-Hatch	0.180	0.093	**0.519**
Terzaghi	0.288	0.104	**0.360**

Table 4.6 Variation statistics of different methods after the sensitivity analysis of the parameters related to the shape of the materials. Table 3 shows each applied parameter.

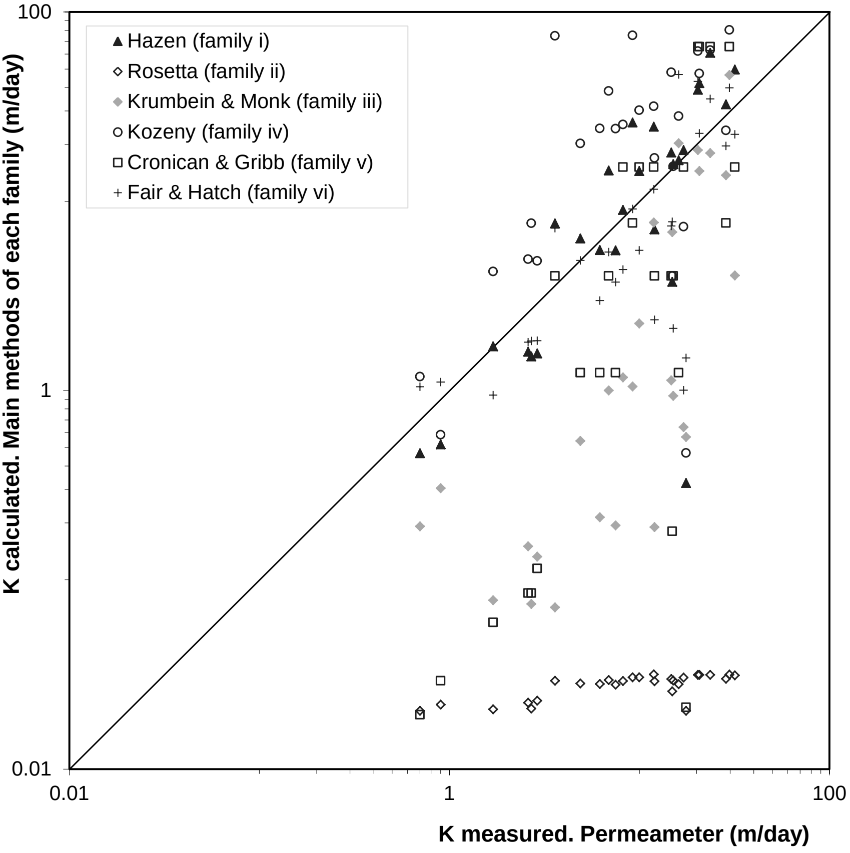


Fig. 4.5 Correlations between the laboratory permeability results (x-axis) and the permeability results obtained by the major methods included in each family (y-axis).

Carman, Rumer, Bakhmeteff, Bakhmeteff-Feodoroff, Sauerbrei, Krüger, Chapuis (2004a), and Chapuis (2004b) methods; and **Family vi)** the Fair-Hatch method.

Upscaling and comparison with the pumping test results

K values from grain size methods are often several times lower than those obtained by pumping tests (Bear, 1972; Cronican and Gribb, 2004;Shepherd, 1989; Uma et al., 1989; Vienken and Dietrich, 2011; Vukovic and Soro, 1992, among others) and our study is not an exception. In fact, the values obtained from methods collected are 2-3 times lower than those obtained with pumping tests. For this reason, an upscaling analysis has been performed comparing its results with those interpreted from pumping tests.

As section 2 describes, samples used in this study belong to the first 8 meters and that the pumping tests in the study area analyse all the thickness of the aquifer (18 meters). The equations to obtain *K* from the lithological parameters represent the static properties at a given scale and, therefore, are not dependent on the geometry and/or boundaries of the aquifer.

According to detailed lithological studies of the study area (Cabelllo et al., 2007) and at a regional scale (Gámez et al., 2009 and the references described therein), the analysed hydrogeological units in this study belong to deltaic prograding structures (mainly flood plains) of the Holocene. This type of structures has a high homogeneity, and on a large scale the grain size of this kind of structures are fining upward. In this context, giving that the upper areas are homogeneously connected porous media, these zones would dominate the flow along the aquifer and, therefore, could be representatives of the effective *K* of the hydrogeological unit. From the mathematically point of view, Sánchez-Vila et al., (2006) argue that high-conductivity zones tend to display better connections than other zones with low-conductivity values, increasing the equivalent hydraulic conductivity values.

Using the power average method 4.3, Journel et al. (1986) suggested that the geometric average is proportional to the equivalent hydraulic conductivity. Thus, methods based on the geometric average should also be proportional to the effective hydraulic conductivity. Figure 4.6 graphically demonstrates that most of the methodologies underestimate *K*, producing values below the lower limit of upscaling (the harmonic average). The Shepherd (1989), Kozeny (1953), Bakhmeteff (Bakhmeteff and Feodoroff, 1937) and Chapuis (2004) methods yields results close to the geometric mean, with a ratio ranging from 0.97 and 1.13.

Based on comparisons with pumping test results, studies have suggested that K values obtained by pumping tests are substantially higher than the geometric average of theoretically predicted point-based K values (Cheong et al., 2008; Dagan, 1989; Sánchez-Vila et al., 1996; Vukovic and Soro, 1992). In this study, the average value from the pumping test (12.4 m/day) is close to the arithmetic mean (12.2 m/day; the upper limit of upscaling). Hence, it is reasonable to assume that the flow is predominantly parallel to the layers or zones in a highly homogeneous connected network. None of the methods that meet the acceptance criteria are close to the geometric mean, except the Slichter (1899) method, with a ratio of 1.00.

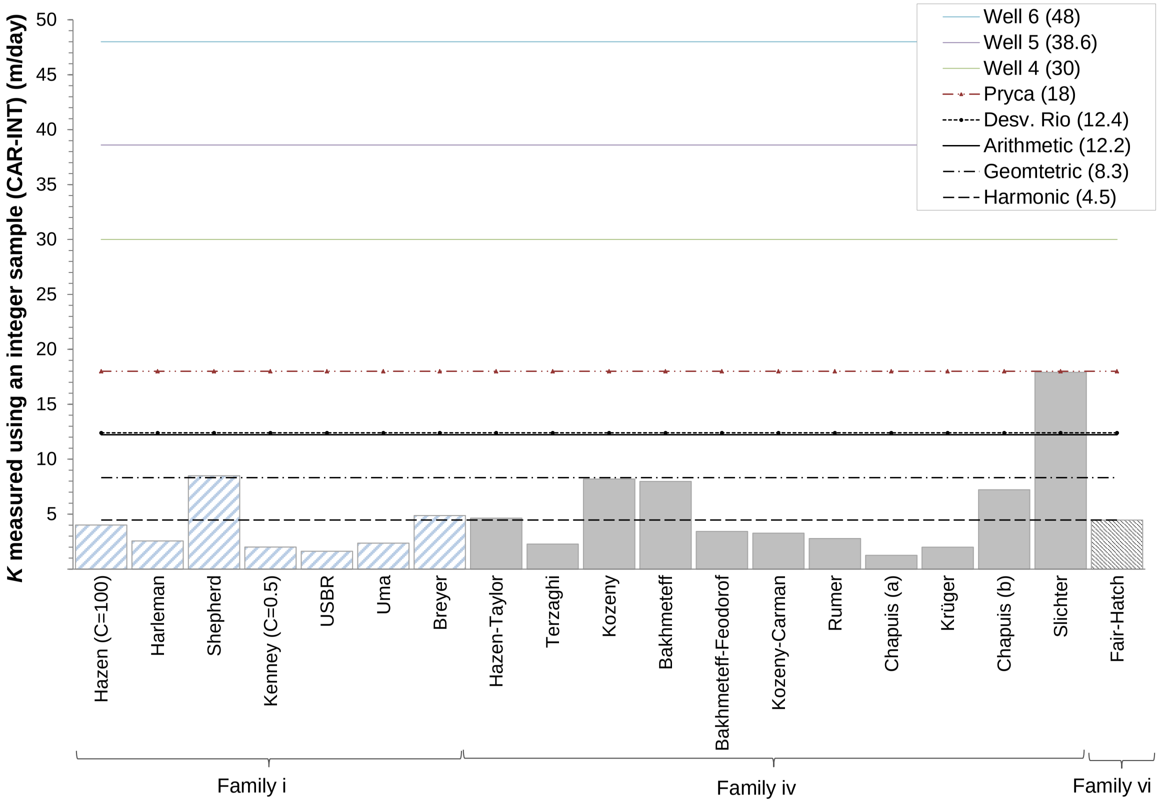


Fig. 4.6 Comparison of the pumping test results and from the formulas that met criteria (1) and (2) where the upscaling method has been applied previously.

4.5 Summary and conclusions

This study addresses the selection of the optimal methodology for estimating the hydraulic conductivity of individual sandy samples at the centimetre scale and then analysing their applicability in an upper scale.

Various methodologies are evaluated for use in upscaling studies. Up to 51 methods of calculating the hydraulic conductivity (K) and intrinsic permeability from the grain size and related parameters were reviewed from the literature. Among them, 41 methods were actually used. The range of application is specified in all the methods. Some of these methods did not perform satisfactorily, likely because this application was beyond the acceptable range. It should also be stressed that some mistakes in the descriptions of certain methods in the literature have been identified and are discussed in this manuscript.

In general, the methods evaluated are highly correlated, but they are not satisfactorily correlated with the measured K. Most of the formulas give K values lower than measured K values, leading to underestimations of K.

Although it is difficult to draw strong conclusions from a statistical analysis with a limited number of samples, we consider the following results.

- The methods of Hazen and Harleman are recommended because of their simplicity and accuracy in providing K estimates.
- The Shepherd method yields good accuracy only if the empirical parameters related to the sample description are properly selected.
- The application of the Bakhmeteff and Bakhmeteff-Feodoroff methods depends on whether the porosity is known. The methods of Kozeny, Terzaghi, Rumer, Sauerbrei, Krüger, Chapuis and Slichter also require empirical constants and display slight oscillations. All of these methods provide reasonable K estimates.
- With respect to the Fair-Hatch method, it should be determined whether the method works equally well with fewer points from the sieve size distribution. However, this method also uses the porosity and other empirical constants.
- Most of the PTFs did not perform well for samples from the test site. The correct equation (or family of equations) should be carefully selected and evaluated before being applied. Most methods in the PTF family do not perform well for samples with little clay, and this grain size is crucial for

> these methods. Thus, these methods are not recommended for sandy samples because the percentage or clay are relevant on their proper functionality.

Upscaling based on the power average method (equation 4.3) was used for the evaluation of limits of the theoretical equivalent hydraulic conductivity. Afterwards, all the equations compiled were applied based on an integrated sample. The evaluation of the application of each formula at larger scale has been performed by comparing their results with those of pumping tests. Many studies have suggested that the geometric mean represents the theoretical equivalent hydraulic conductivity of a given area. Shepherd (1989), Kozeny (1953), Bakhmeteff (Bakhmeteff and Feodorof, 1937) and Chapuis (2004) supports these studies. Additionally, The Slichter method (1899) results are close to those obtained with pumping tests. Thus, it may suggests its application to obtain an equivalent hydraulic conductivity that fits with the results obtained with field tests, which have a greater scale of observation. These first results could be a step forward in the application of these equations at lager scale by applying them in an integrated sample. For instance, this methodology can be applied as a first step in the hydraulic parameterization of hydrogeological studies in an inexpensive manner.

Finally, the conclusions and recommendations presented can be applied to sandy samples. Further studies using other grain sizes and other hydrogeological contexts are required to determine the most suitable method(s) for each type of these contexts.

Chapter 5

AkvaGIS: An open source tool for water quantity and quality management

Summary

AkvaGIS is a novel, free and open source module included in the FREEWAT plugin for QGIS that supplies a standardised and easy-to-use workflow for the storage, management, visualisation and analysis of hydrochemical and hydrogeological data. The main application is devised to simplify the characterisation of groundwater bodies for the purpose of building rigorous and data-based environmental conceptual models (as required in Europe by the Water Framework Directive). For data-based groundwater management, AkvaGIS can be used to prepare input files for most groundwater flow numerical models in all of the available formats in QGIS. AkvaGIS

Based on: Criollo, R., Velasco, V., Nardi, A., Manuel de Vries, L., Riera, C., Scheiber, L., Jurado, A., Brouyère, S., Pujades, E., Rossetto, R., Vázquez-Suñé, E., 2018. AkvaGIS: An open source tool for water quantity and quality management. Computers & Geosciences. doi:10.1016/J.CAGEO.2018.10.012.

is applied in the Walloon Region (Belgium) to demonstrate its functionalities. The results support a better understanding of the hydrochemical relationship among aquifers in the region and can be used as a baseline for the development of new analyses, *e.g.*, further delineation of nitrate vulnerable zones and management of the monitoring network to control chemical spatial and temporal evolution. AkvaGIS can be expanded and adapted for further environmental applications as the FREEWAT community grows.

5.1 Introduction

Environmental assessment and characterisation of groundwater bodies (as required by the Water Framework Directive; European Commission, 2000) involve continuous monitoring and evaluation of a large number of physical and chemical parameters (*e.g.*, groundwater level, temperature, pH, or nitrates, among others). These parameters, which are used to conceptualise the behaviour of the environmental system, can be reinforced by other information (such as geology or isotopes) and are often stored in different scales and formats (*e.g.*, maps, spreadsheets or databases). This conceptualisation of the environment is essential for the development of numerical models (Refsgaard et al., 2010), which are common and effective tools used to obtain deeper insights into physical systems. For instance, groundwater numerical models supported by hydrochemical data are used to (i) control different flow paths and their relationships among different water bodies, (ii) characterise water-rock interactions, (iii) identify water quality spatial and temporal evolutions, (iv) evaluate groundwater storage changes, and (v) design strategies to achieve a good chemical status based on national/international thresholds for water quality, among others. With respect to the latter, water agencies, stakeholders and water suppliers usually encounter difficulties in ensuring compliance with standard regulatory guidelines (Gleeson et al., 2012; Jurado et al., 2017; Vázquez-Suñé et al., 2006; 2016a).

Geographical Information Systems (GIS) supply useful tools for addressing the above mentioned issues in collection, archiving, analysis, and visualisation of spatial and non-spatial data in different formats. GIS software is widely used by the scientific community, public administration and the private sector. The comprehensive application of GIS platforms can aid in producing environmental assessments such as evaluation of water quality, water availability, zone mapping and risk assessment from the local to regional scale (Duarte et al., 2018; Ghosh et al., 2015; Tiwari et al.,

2017) and improving numerical modelling processes (Kresic and Mikszewski, 2012; Rios et al., 2013; Rossetto et al., 2013; Steyaert and Goodchild, 1994), among other applications.

Several authors have developed GIS techniques within licensed GIS platforms to optimise environmental analyses (*e.g.*, Chenini and Ben Mammou, 2010; Kim et al., 2012; Lee et al., 2018) and address groundwater quality issues (*e.g.*, Ashraf et al., 2011; Babiker et al., 2007; Marchant et al., 2013; Nas and Berktay, 2010). These broadly applied advancements were mostly developed in commercial GIS platforms, the commercial licence of which is an obstacle for communities/institutions with limited resources, and these entities are consequently unable to benefit from this technology. Additionally, certain of these developments are not open source, and thus they cannot be expanded and/or adapted for tailored or further applications by third parties. However, these efforts have approached common conceptual and technical issues through creation of GIS-based tools related to (i) management and integration of a notably large amount of time-dependent and spatially dependent data (Cabalska et al., 2005; Chesnaux et al., 2011; Gogu et al., 2001; Maidment, 2002; Strassberg, 2005; Velasco, 2013; Wojda et al., 2006); (ii) homogenization and harmonization of data collected from diverse sources obtained with different techniques (de Dreuzy et al., 2006; Lètourneau et al., 2011; Romanelli, 2012); (iii) communication and data exchange in different formats (Kingdon et al., 2016; Wojda and Brouyère, 2013); (iv) management of hydrological, geological, hydrogeological and hydrochemical data with diverse temporal and spatial ranges (Criollo et al., 2016; Merwade et al., 2008; Vázquez-Suñé et al., 2016a; Velasco, 2013; Velasco et al., 2014); and (v) analysis of the required spatio-temporal data oriented to pre- and post-processing and generation of hydrogeological models (Alcaraz et al., 2017; Li et al., 2016; Strassberg et al., 2011; Wang et al., 2016).

Given these obstacles, the need becomes clear for open-source and user-friendly software that allows free access to the groundwater community for both application and further developments to adapt these tools to specific institutions and/or third-party databases (Bhatt et al., 2008; Dile et al., 2016). Selected open-source GIS-based tools are available that address these requirements for other topics, such as aquatic ecosystems assessments (Nielsen et al., 2017), which are beyond the scope of the current study but can be found in Khosrow et al. (2013), Ye et al. (2013), Teodoro (2018) or Huang et al. (2018). For groundwater management, open-source and GIS-based tools designed without specific user-friendly tools for hydrochemical and hydrogeological analyses in a unique GIS platform were developed to homogenise, integrate and visualise groundwater-related data (Boisvert et al. 2007, 2012; Jarar Oulidi et al. 2009, 2015) and to connect GIS platforms with groundwater numerical models (Bhatt et al., 2008, 2014; Carrera-Hernández et al., 2008; Rossetto et al., 2013). Hence, new open-source GIS-based software should allow standardization, management, analysis, interpretation and sharing of hydrogeological and hydrochemical data within a unique geographical context.

To address all of the aforementioned issues, a unique free and open-source GIS-integrated environment for water resource management with special reference to groundwater was developed in the context of the H2020 FREEWAT project (freewat.eu). The main objective was to promote the application of EU (WFD; European Commission, 2000) and other water-related directives (De Filippis et al., 2017; Foglia et al., 2018; Rossetto et al., 2018). FREEWAT is a large QGIS plugin (QGIS Development Team, 2009) (Figure 5.1) in which all data related to surface and subsurface water bodies can be digitised, archived, analysed (also using integrated numerical models) and visualised. Additionally, the FREEWAT concept aimed to perform extensive capacity-building activities in an innovative participatory approach by gathering technical staff and relevant stakeholders for proper application of water policies (Criollo et al., 2018a; De Filippis et al., 2018; Foglia et al., 2017).

In this chapter, we present the AkvaGIS tool, a user-friendly, free and open-source GIS-based package integrated into the FREEWAT platform (Figure 5.1). AkvaGIS has been designed to fulfil the needs for: (i) managing and visualising hydrogeological and hydrochemical standardised data with different temporal and spatial scales to facilitate development of the environmental conceptual model, (ii) integrating data from diverse sources gathered using different data access techniques and formats, and (iii) preparing hydrogeological input files for any groundwater numerical model in all of the available formats in QGIS. Due to its open-source architecture, AkvaGIS can be updated and extended by any advanced user.

After a description of the AkvaGIS design and relevant tools in the following section, we present an application in the Walloon region (Belgium) to demonstrate certain capabilities. Finally, the development, the application and future improvements are discussed.

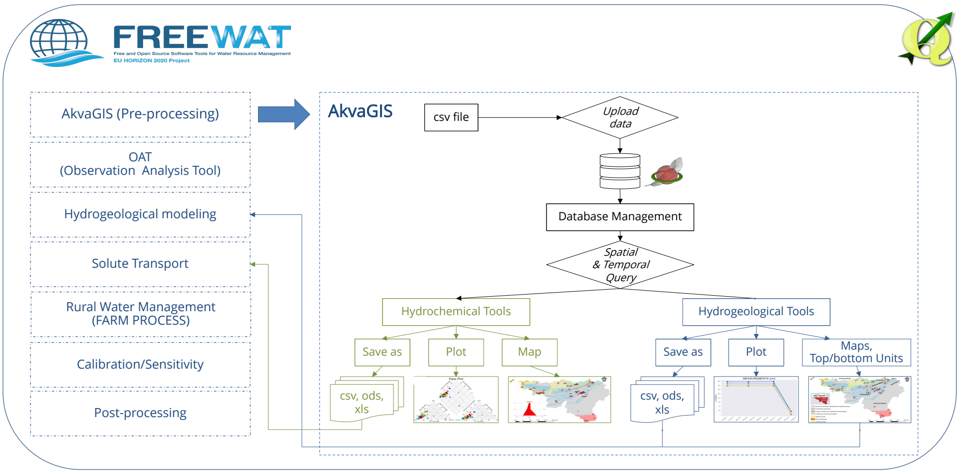


Fig. 5.1 AkvaGIS tools: Scheme of simplified workflow together with all FREEWAT tools. Colours correspond to the 3 main groups of tools: database management (black), hydrochemical tools (green) and hydrogeological tools (blue).

5.2 AkvaGIS description

AkvaGIS is the evolution of work performed by Velasco (2013), Velasco et al., (2014), Alcaraz (2016) and Criollo et al., (2016). In those studies, tools for geological, hydrochemical, geothermal and hydrogeological data analysis were developed in the commercial GIS desktop software ArcGIS (ESRI, 2005, 2012). Conversely, AkvaGIS is a free and open-source GIS-based tool supported in Linux and Windows OS and integrated in QGIS (Criollo et al., 2017). QGIS is supported by most operating systems (Windows, Linux, Unix, Macintosh) and has several data reading and writing formats. The data management subsystems allow easy and rapid queries that are quickly processed and displayed, and this tool has a large community of developers (Chen et al., 2010; Bhatt et al., 2014).

5.3 Software design and structure

AkvaGIS is developed in Python (python.org) and integrated into the FREEWAT platform (Figure 5.1). This tool is freely available from the official QGIS experimental repository, the FREEWAT project repository (freewat.eu) or the gitlab repository (gitlab.com/freewat). The AkvaGIS tools enhance FREEWAT with hydrochemical and hydrogeological data processing and analysis. AkvaGIS is designed to avoid code repetition to reduce errors and improve the code maintenance under the GNU Lesser General Public License v2.0 (GPL, GNU) or later. Different third-party libraries are applied with GPL, MIT (MIT) and BSD (BSD) license types. The Python-related dependencies that AkvaGIS applies are the Qt version 4 Python wrapper (PyQt4), a Python 2D plotting library that creates quality figures in a variety of hardcopy formats and interactive environments across platforms (Matplotlib 1.5, ChemPlotLib 1.0, Openpyxl 2.3, Odfpy 1.3. and Pyexcel 0.2). All of these libraries are automatically downloaded during FREEWAT installation.

AkvaGIS tools are divided into 3 main sections (Figure 5.2): (1) the database management tools that are designed to manipulate the hydrochemical and hydrogeological data stored in the AkvaGIS database; (2) the hydrochemical tools for managing, visualising, analysing, interpreting and pre-processing the hydrochemical data and; (3) the hydrogeological tool. This package was developed to facilitate interpretation of hydrogeological information and hydrogeological units, which in turn is crucial in defining conceptual models and in modelling activities. The hydrochemical and hydrogeological tools allow creation of contour maps and further spatial operations. Additionally, thematic maps (*e.g.*, chlorides, piezometric maps or pumping rates) can be created for the selected points and time periods using different functionalities included in the AkvaGIS menu.

AkvaGIS database

The core of the AkvaGIS tools is a geospatial database (Figure 5.3) implemented using the relational database SpatiaLite (SQLite spatial extension, sqlite.org), where all data related to a hydrogeological study are stored. SpatiaLite is an open-source database able to store many format files (*e.g.*, raster, shapefiles or cad files), and it can build-in spatial indices, which facilitate rapid searches over large areas. A SpatiaLite database can be safely exchanged across different platforms because its internal architecture is universally portable (Spatialite Development Team, 2011). Accordingly, this database can be expanded and/or adapted for future applications and can be continuously improved. No-installation and no-configuration are required before using the AkvaGIS database file within QGIS.

The AkvaGIS database architecture can store a large amount of spatial features and hydrochemical and hydrogeological temporal-dependent data and is designed for different methodologies and tools used by water professionals and managers to

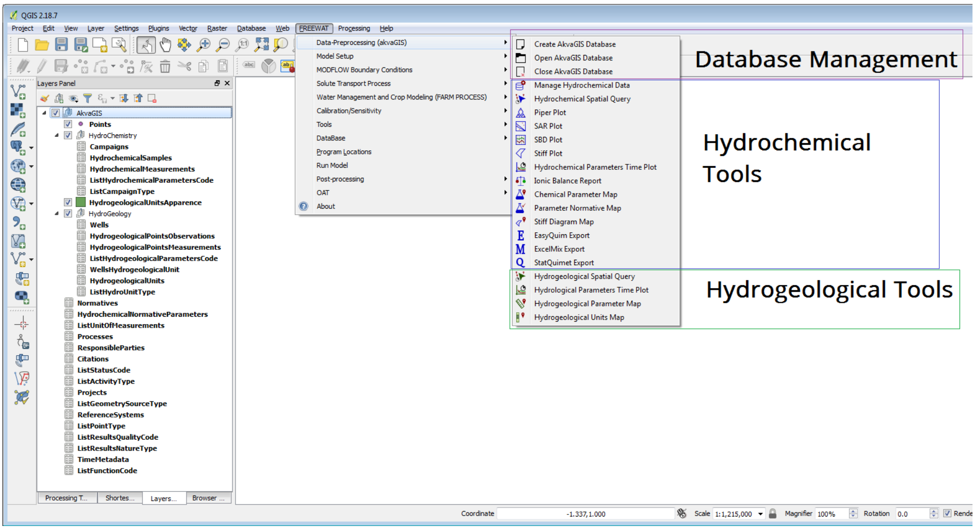


Fig. 5.2 FREEWAT menu of tools, including AkvaGIS tools, are presented in the QGIS layer panel (version 2.18 Las Palmas). AkvaGIS menu shows the three groups of tools: database management (black), hydrochemical (green) and hydrogeological (blue) analyses.

address groundwater management issues. AkvaGIS considers the aforementioned existing projects (*e.g.*, Strassberg, 2005; Wojda and Brouyere, 2013) and implements selected international standards to store and exploit hydrogeological information, time series, and field observations and measurements. These standards are supported by the Open Geospatial Consortium (OGC, 2003; 2006; 2007; OGC Water ML 2.0, 2012), GeoSciML (Sen and Duffy, 2005), the specifications of the European Directive INSPIRE (2011; 2013) and the ONEGeology project (2013). Hence, the standardised architecture facilitates harmonization of the collected data, and the AkvaGIS database can be shared in a more understandable manner.

The spatial coordinates of the points (*i.e.*, piezometers, wells, springs, swallow holes, seeps, vanishing points or any other specific points from water bodies) related to the location of measurements/estimates or collected samples are the basic information required for use of the AkvaGIS tools and are stored in the *Points* table. The basic hydrochemical information related to each spatial point, *i.e.*, *HydrochemicalSamples* and *HydrochemicalMeasurements* tables (Figure 5.3), contains the dates when each named sample was collected, the dates of the physical and chemical parameters analysis, and their corresponding values and units. The list of analysed parameters is stored in a library/catalogue (*ListHydrochemicalParametersCode*) and can be updated by the user.

Similarly, the basic hydrogeological information is related to the corresponding spatial point at which the hydrogeological measures/estimates were collected. The measurement dates, measurements and estimated parameters and the corresponding values and units are stored in the tables *HydrogeologicalPointsObservations* and *HydrogeologicalPointsMeasurements.* The default hydrogeological parameters available in the library/catalogue *ListHydrogeologicalParametersCode* store flow rate, depth to water, pressure and hydraulic head. This list of parameters can be customised by the user. The hydrogeological unit observed at each point can be defined and stored in the tables *HydrogeologicalUnits* and *WellsHydrogeologicalUnit* (see Figure 5.3). These interpreted units can be interpolated to generate surfaces of the boundaries of hydrogeological units of the study zone, which might be subsequently applied to define the three-dimensional geometry in groundwater numerical models. The created files can be saved in any format available in QGIS such as shapefile or raster.

Additional information can also be stored, such as field campaign number, entities in charge of measurements or responsible parties, among others. This information is not essential for use of the AkvaGIS tools, but it is useful in managing the hydrogeological and hydrochemical data. Detailed information on all AkvaGIS tables and their fields are shown in the FREEWAT user manual volume 4 (Serrano-Juan et al., 2017).

Through the QGIS project (.qgis file), the user manages the AkvaGIS database (.sqlite file) and additional files that are shown in the layer panel (see Figure 5.2). The Database Management tools allow the user to create, open and close the AkvaGIS database. Once the information collected is stored in the AkvaGIS database, users can apply the analysis tools (Figure 5.3).

Hydrochemical analysis tools

The Hydrochemical Analysis Tools package supplies a wide range of tools for performing hydrochemical data analysis through common queries and hydrochemical plots. The ***"Manage Hydrochemical Data"*** tool allows visualisation of hydrochemical data from points already stored in the AkvaGIS database. The user can manage these data by adding, deleting or editing the needed information to perform the study (see Figure 5.4a).

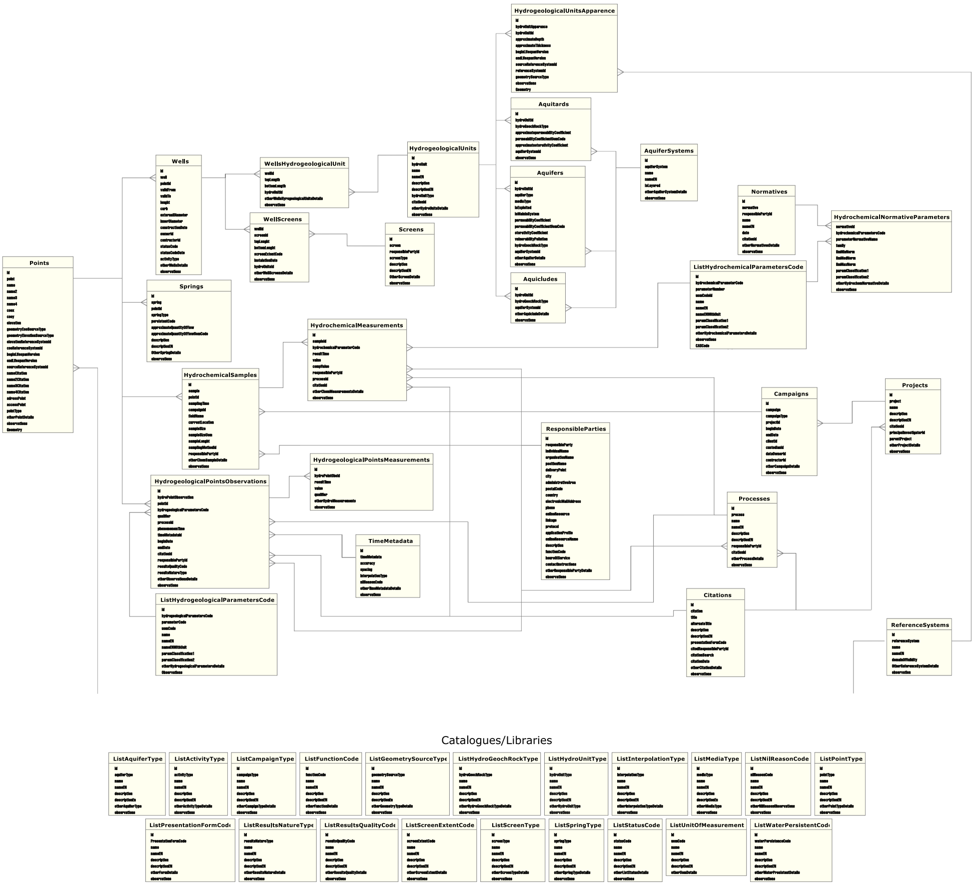
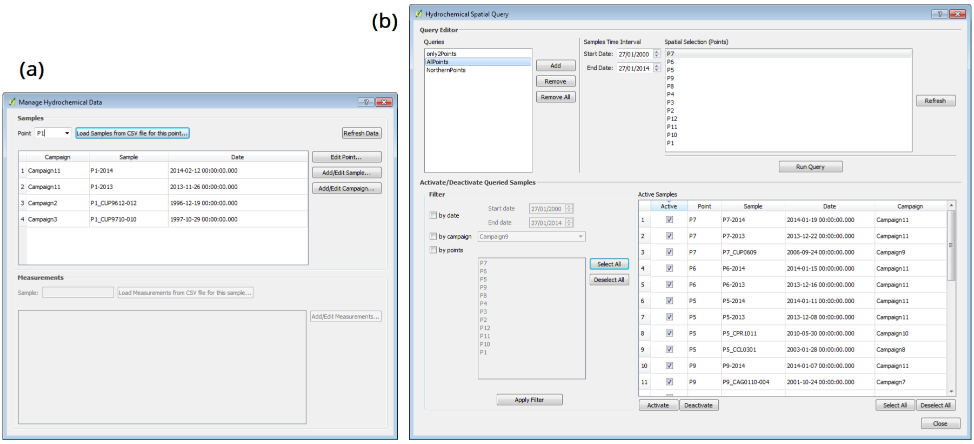


Fig. 5.3 AkvaGIS geospatial database scheme.

First, a selection query must be run to create diagrams and maps. This query is created and stored in the database for future application. The ***"Hydrochemical Spatial Query"*** tool performs a specific selection of points in the desired time period (see Figure 5.4b). Selection using reference campaigns, dates or geographical position can be performed. Diagrams and maps preparation use the queries initiated with this tool. The query results can be saved in a table for further external analysis.



Point: P1

	Campaign	Sample	Date
1	Campaign11	P1-2014	2014-02-12 00:00:00.000
2	Campaign11	P1-2013	2013-11-26 00:00:00.000
3	Campaign2	P1_CUP9612-012	1996-12-19 00:00:00.000
4	Campaign3	P1_CUP9710-010	1997-10-29 00:00:00.000

Queries: only2Points, AllPoints, NorthernPoints

Start Date: 27/01/2000

End Date: 27/01/2014

Spatial Selection (Points): P7, P6, P5, P9, P8, P4, P3, P2, P12, P11, P10, P1

☐ by date Start date 27/01/2000 End date 27/01/2014

☐ by campaign Campaign9

☐ by points P7, P6, P5, P9, P8, P4, P3, P2, P12, P11, P10, P1

	Active	Point	Sample	Date	Campaign
1	☑	P7	P7-2014	2014-01-19 00:00:00.000	Campaign11
2	☑	P7	P7-2013	2013-12-22 00:00:00.000	Campaign11
3	☑	P7	P7_CUP0609	2006-09-24 00:00:00.000	Campaign9
4	☑	P6	P6-2014	2014-01-15 00:00:00.000	Campaign11
5	☑	P6	P6-2013	2013-12-16 00:00:00.000	Campaign11
6	☑	P5	P5-2014	2014-01-11 00:00:00.000	Campaign11
7	☑	P5	P5-2013	2013-12-08 00:00:00.000	Campaign11
8	☑	P5	P5_CPR1011	2010-05-30 00:00:00.000	Campaign10
9	☑	P5	P5_CCL0301	2003-01-28 00:00:00.000	Campaign8
10	☑	P9	P9-2014	2014-01-07 00:00:00.000	Campaign11
11	☑	P9	P9_CAG0110-004	2001-10-24 00:00:00.000	Campaign7

Fig. 5.4 Using Manage Hydrochemical Data (5.4a), tool users can update, upload or delete data stored in the AkvaGIS database. Diagram and maps are created by applying a Hydrochemical Spatial Query (5.4b), which is subsequently stored in the same database for further analysis.

The ***"Ionic Balance Report"*** tool allows calculation of the ionic balance report (shown in Figure 5.5a). This tool automatically converts all units to *meq/l* and selects the major ions of the chosen sample. Once the query is created, the user can save the results in a table or in an ionic balance report (.ods format).

AkvaGIS offers the ability to draw a number of hydrochemical diagrams useful for analysing the water chemical composition and how the collected samples relate to each other. The ***"Piper Plot"*** is useful for visualising hydrochemical types of water samples classified by their ionic composition. The ***"SAR Plot"*** (Sodium Adsorption Ratio diagram, Figure 5.5b) is useful for analysis of irrigation water quality to facilitate the management of sodium-affected soils. To visualise and analyse water mixing, end-members or changes between certain ionic relationships, users can apply the ***"Shöeller-Berkaloff diagrams"***. With the ***"Stiff Plot"***, the user can analyse the samples compositions in its spatial context among water from different sources. Figure 5.5c presents the interfaces developed to manage these diagrams. Plot setup commands (plot size, point style, legend, among other configurations) are available in the AkvaGIS diagram and map tools.

Spatial analysis is useful in visualising and analysing the hydrochemical spatial variation throughout the study zone. To this end, the ***"Chemical Parameter Map"*** and the ***"Stiff Diagram Map"*** tool supply spatial distribution analysis of the chemical samples and the Stiff diagram zone, respectively.

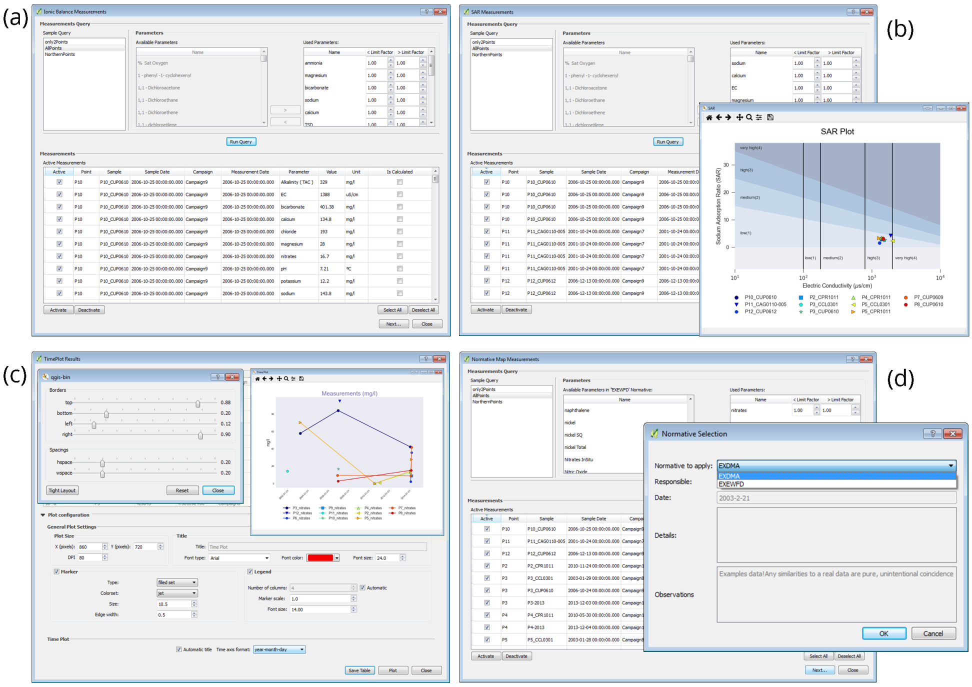
The temporal distribution of chemical parameters can be analysed by drawing a ***"Time Plot"*** of the query previously created using the *Hydrochemical Spatial Query* tool (Figure 5.5c). In addition, tools are available to export the query data to different external platforms for evaluation of common major ions (*e.g.*, Easy_Quim; Serrano-Juan et al., 2015), mixing ratios of the samples (*e.g.*, MIX; Carrera et al., 2004) or ionic relationships and further statistical analyses (*e.g.*, Statistical Tools; Velasco 2013). These tools are available at the Hydrogeology Group (UPC-CSIC) website, where other relevant tools can be downloaded for further groundwater analysis (https://h2ogeo.upc.edu/).

The ***"Parameter Normative Map"*** draws thematic maps according to the threshold values for the queried parameters established by a given guideline (*e.g.*, Water Framework Directive) (**Figure 5d**). The guideline and their thresholds values must be uploaded and stored previously in the database by the user.

Hydrogeological analysis tools

This module presents a set of tools developed to improve management, visualisation and interpretation of the hydrogeological data. The user can manage and query hydrogeological measurements and estimates performed in wells, piezometers or springs. Thematic maps of each chosen parameter (*e.g.*, piezometric maps) can be performed from the selected points and the specific time interval. General statistics can be calculated for each selected parameter to perform simple analyses of the temporal data. Additionally, this tool can simplify the construction of the geometry of groundwater flow numerical models. Hence, these tools create depth or thickness surfaces of the defined hydrogeological units (top and bottom of each layer). The user can save these structures in several formats with the QGIS tools and apply them in a groundwater numerical model (*e.g.*, MODFLOW).

Similarly, the ***"Hydrogeological Spatial Query"*** tool enables consultation of the hydrogeological measurements (*i.e.*, head level, water depth, pumping rates and



Normative to apply: EXDMA

Date: 2003-2-21

Examples data! Any similarities to a real data are pure, unintentional coincidence

Fig. 5.5 Examples: Ionic Balance Measurements (5a); Sodium Adsorption Ratio diagram (5b); Time Evolution Plot (5c) and Normative Maps (5d).

discharge) collected in wells, piezometers or springs. This query only acts on those points where hydrogeological observations and measurements have been introduced in the database. This command creates and adds spatial queries of the selected points (spatial selection) for the desired time interval. Different methods are used to create this selection: by sampling campaigns, by dates or by the geographical positions. The interface uses the same commands as the hydrochemical spatial query interface.

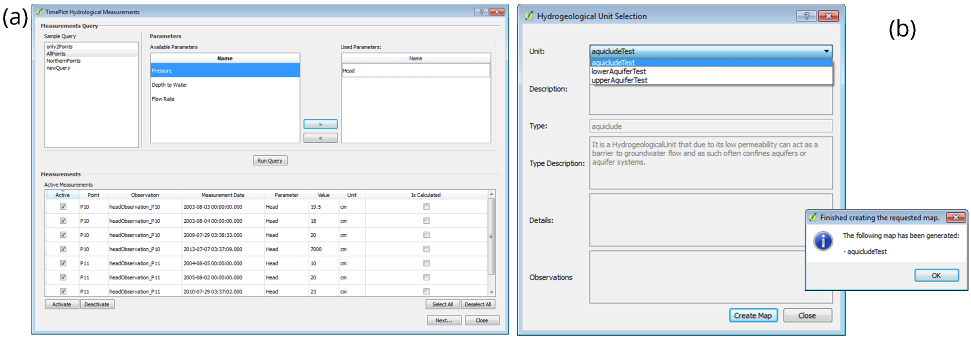
In selecting the previously created desired hydrogeological spatial query, the user can create a time evolution plot of the chosen parameters using the ***"Hydrogeological Parameter Time Plot"*** tool (shown in Figure 5.6a). Additionally, the ***"Hydrogeological Parameter Map"*** tool creates parameter maps of the selected query for the desired parameters. The available hydrogeological parameters are depth to water, flow rate, head or pressure (as listed in the library/catalogue *ListHydrogeologicalParametersCode*). The user is able to choose the value used in the map (earliest, latest, minimum, maximum or average) to draw the most important information.

Three-dimensional groundwater flow numerical models require the definition of aquifer geometry. Therefore, a modeller must build surfaces limiting different hydrogeological units, as defined in the conceptual model. The ***"Hydrogeological Units Maps"*** tool (shown in Figure 5.6b) creates maps of top/bottom hydrogeological units. Because the FREEWAT plugin (Rossetto et al., 2018) includes MODFLOW (Harbaugh, 2005) as numerical code for groundwater flow simulation, the user can save these geometrical boundaries in a proper format for later implementation in a numerical model working in the same GIS environment.

For all tools described above, the results can be saved as tables, and the corresponding plots and maps can be user-customised. FREEWAT user manual volume 4 (Serrano-Juan et al., 2017) and the training material contain additional information on the AkvaGIS functionalities.

5.4 An example of AkvaGIS application: The Walloon Region (Belgium)

Database created for this example can be found in Criollo et al. (2018b) and it can be free downloaded to reproduce this study.



Active	Point	Observation	Measurement Date	Parameter	Value	Unit	Is Calculated
☑	P10	headObservation_P10	2003-08-03 00:00:00.000	Head	19.5	cm	☐
☑	P10	headObservation_P10	2003-08-04 00:00:00.000	Head	18	cm	☐
☑	P10	headObservation_P10	2009-07-29 03:38:33.000	Head	20	cm	☐
☑	P10	headObservation_P10	2013-07-07 03:37:09.000	Head	7000	cm	☐
☑	P11	headObservation_P11	2004-08-05 00:00:00.000	Head	10	cm	☐
☑	P11	headObservation_P11	2005-08-02 00:00:00.000	Head	20	cm	☐
☑	P11	headObservation_P11	2010-07-29 03:37:02.000	Head	23	cm	☐

Unit: aquicludeTest

Type: aquiclude

Type Description: It is a HydrogeologicalUnit that due to its low permeability can act as a barrier to groundwater flow and as such often confines aquifers or aquifer systems.

Finished creating the requested map.

The following map has been generated:

- aquicludeTest

Fig. 5.6 Interface of the Time Plot Hydrogeological Measurements (6a) and Hydrogeological Unit maps (6b).

Thus far, AkvaGIS and all of the FREEWAT tools have been extensively used by more than 1300 attendees during courses held in over 50 countries. The AkvaGIS tools have been further improved and developed to facilitate its handling because of the feedback supplied by these users.

In the following text, we present an application of selected AkvaGIS tools with real data to demonstrate the advantages of use. Specifically, AkvaGIS is applied to data collected in the Walloon region (southern region of Belgium-northwest Europe; Figure 5.7). The Walloon region has an area of approximately 16,844 km^2, where half of the land is covered by agricultural areas and forests (approximately 30%) and urban areas (approximately 15%) (Brahy, 2014). The Walloon region can be roughly divided into six main aquifer units characterised by geological age. Most aquifers are located in fractured rock systems that show a high degree of heterogeneity (Figure 5.7). These aquifers can be distinguished by consolidation: (1) unconsolidated aquifers where groundwater is stored in the interstices of the subsoil (*e.g.*, Tertiary sands and Quaternarian deposits, Figure 5.7) and (2) consolidated aquifers where groundwater is abstracted from permeable and fractured areas (*e.g.*, Primary limestones, Figure 5.7) (SPW-DGO3, 2016).

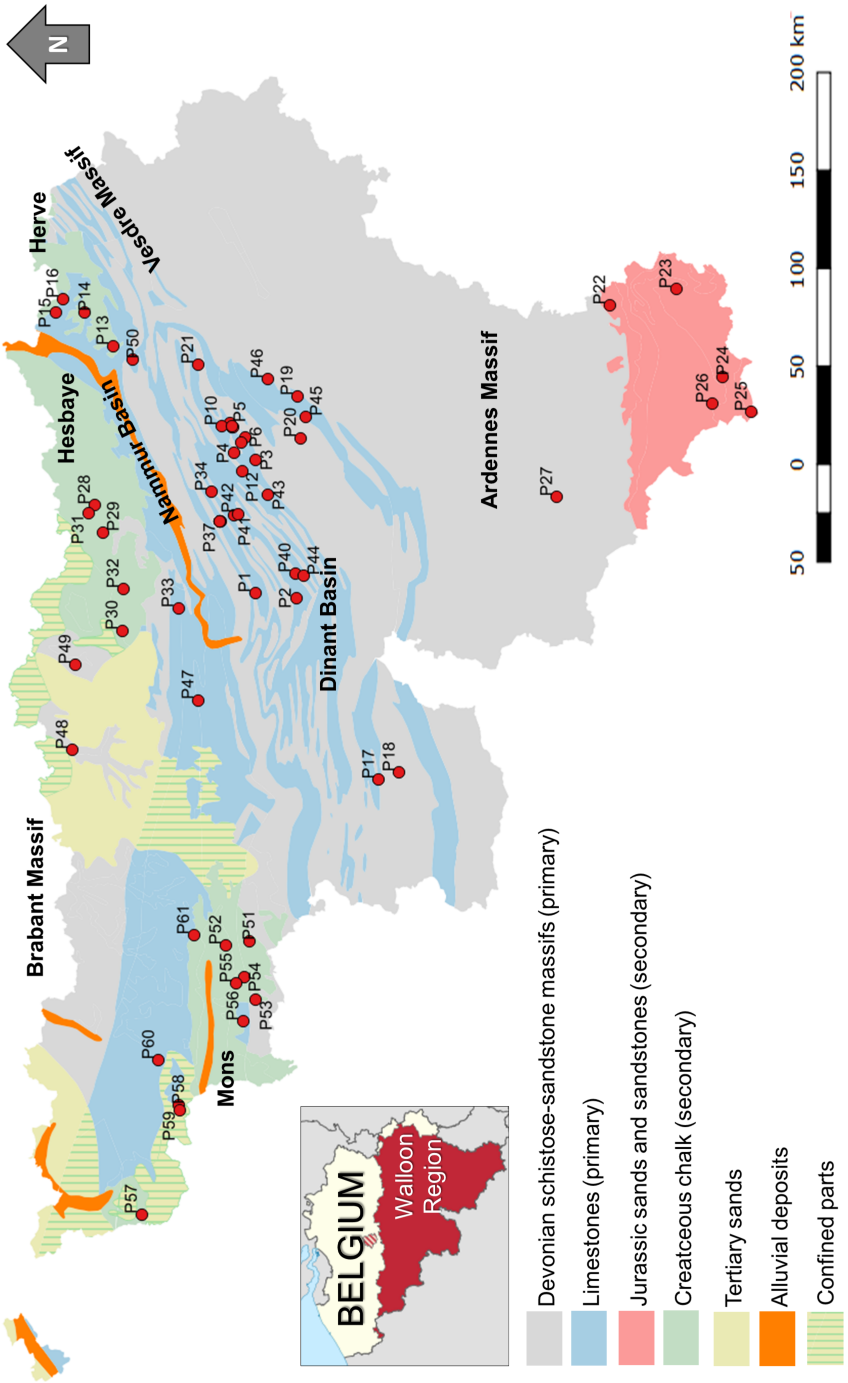


Fig. 5.7 Location of the study area (Walloon Region, Belgium) with the main aquifers. Sampling points are shown as red points.

The majority of groundwater abstraction originates from limestone aquifers (51%, in blue) and chalky formations (21%) and is mainly applied for water supply purposes, representing up to 80% of the water volumes collected ($400 \cdot 10^6$ m^3/y; SPW-DGO3, 2016).

A total of 64 groundwater samples were collected in spring 2016 within the framework of a project that investigated the occurrence and indirect emissions of greenhouse gases (GHGs) from groundwater at the regional scale (Jurado et al., 2018). Analysis of these samples included GHGs, major and minor ions and metals. Data from major and minor ions are used in this paper to display the functionalities of AkvaGIS. Database created for this purpose can be found in Criollo et al. (2018b).

After collecting and storing all of the data in the AkvaGIS database, the chemical analysis quality for charge balance was calculated. In total, 94% of the samples had less than ±5% error (considered acceptable for this study). Once the hydrochemical data quality was ensured, the second step analysed the hydrogeochemical data using graphical diagrams. AkvaGIS generated a map presenting the Stiff plot for each sample. A quick review of this map shows that most of the groundwater samples could be classified as Ca-HCO_3 types (see Figure 5.8).

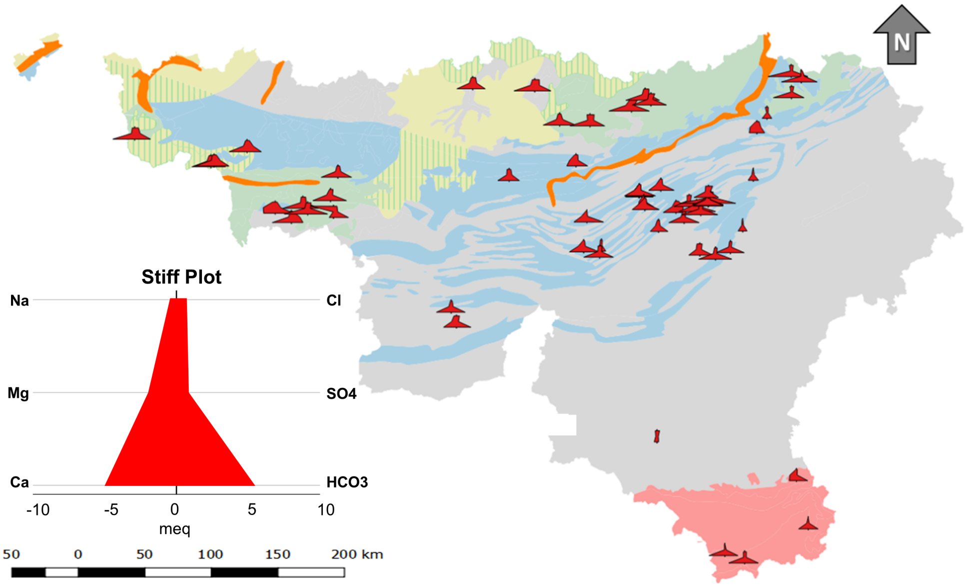


Fig. 5.8 Spatial distribution of the Stiff diagram of the Walloon region aquifers as generated from data collected in the 2016 campaign.

These observations can be corroborated by generating Piper and Schoeller-Berkalof plots (Figure 5.9). Although these plots do not provide supply information on spatial distribution, they allow identification of the main trends with respect to the chemical composition of water samples. The plots also show that most samples have a similar composition (Ca-HCO_3 type). However, it is possible to identify two samples with differing compositions, *i.e.*, samples S51 and S27. Sample S51 has a high Na-HCO_3 concentration, and sample S27 is less mineralised but richer in potassium than the remainder of the samples. Note that the S24 sample (Jurassic sands and sandstones) is completely opposite of the previously described sample (Figure 5.9b), with the lowest value of sodium and chlorides. The information derived from these plots is highly useful in defining the characteristics of aquifers (groundwater and rock chemical compositions should be related), residence times (the degree of mineralisation might be related to the residence time) and/or potential uses (*e.g.*, water with high concentrations of Na^+ and low of Ca^{2+} is not advisable for irrigation purposes because it tends to reduce the permeability of the soil, IGME, 2002).

AkvaGIS tools also produce distribution maps for the nitrate concentration measured at different points. The nitrate concentrations show a strong spatial variability (see Figure 5.10), especially in locations close to agricultural and farm areas. Furthermore, according to the Drinking Water Directive (UE Comission, 1998, stored in the AkvaGIS database), the nitrate concentrations of 16% of the samples exceed the threshold value of 50 mg/l. These points are located in the Chalk zone, the most mineralised aquifer (908.6 μS/cm in the Chalk aquifer of the Hesbaye and 844.8 μS/cm in the Chalk aquifer of the Mons). Conversely, the Devonian limestones (shales and sandstones) of the Dinant Basin presented lower values of electrical conductivity, less than 430 μS/cm. The average dissolved oxygen concentrations showed that the groundwater had oxic conditions, ranging from 4 mg/L to 9.1 mg/L (see Table 5.1). Finally, the average temperatures presented little variation, varying from 10.2 °C to 13.4 °C.

The application of the AkvaGIS tools in the Walloon Region case study helped to: (i) visualise and analyse data easily and quickly and (ii) improve understanding of the hydrochemical relationships among aquifers in the region. These initial results might aid water resource authorities in design of future management and monitoring strategies to continue preservation of the quality and quantity of groundwater resources. For example, vulnerable zones due to high nitrate concentrations could be further delineated and the current monitoring network could be managed to control their spatial and temporal evolution using the AkvaGIS tools. The presented analysis

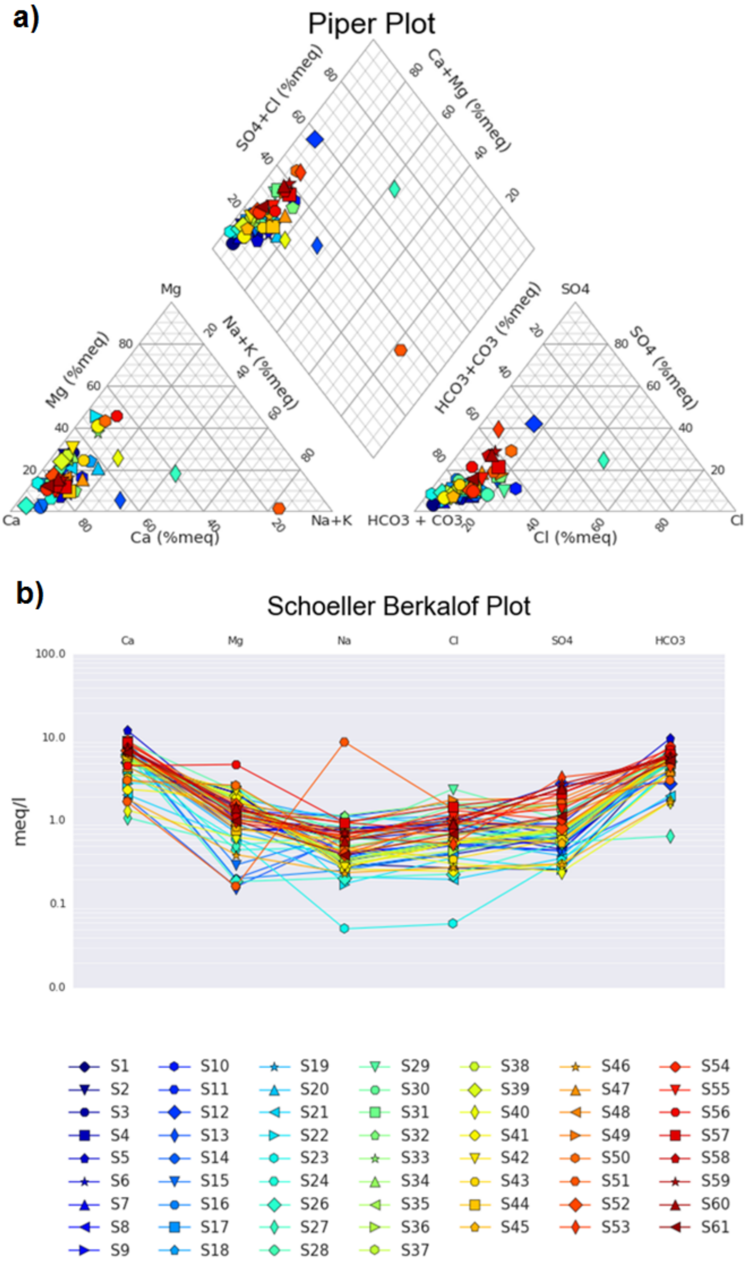


Fig. 5.9 Piper (a) and Schoeller-Berkaloff (b) diagrams of the sampling points from the spring 2016 campaign. Note that S27 and S51 samples have a stronger deviation with respect to the remainder of the samples.

Aquifer formation	Aquifer ID	pH	DO (mg/L)	EC (µs/cm)	T^a (^oC)
Devonian schisto-sandstone massifs (shales and sandstones)	Ardenne Massif	6.74	6	560	12.2
	Dinant Basin	7.59	6.4	552.5	10.8
Primary limestones	Namur Basin	7.19	4	788.7	13.4
	Dev. Dinant Basin	7.58	8.1	425.5	10.2
	Carb. Dinant Basin	7.21	6.4	732.9	11
Jurassic formations (sands and sandstones)	Formations Sud Luxembourg	7.51	4.7	521.6	10.5
Cretaceous chalks	Chalks of Mons	7.13	8.8	844.8	12.8
	Chalks of Hesbaye	7.49	8.4	908.6	13.2
	Chalks of Pays de Hervé	7.03	6.9	671.8	10.5
Tertiary sands	Bruxellian and Landenian Sands	7.37	9.1	736	11.1

Table 5.1 Average of the in-situ parameters of each aquifer.

can be extended to other regions for the same or other water analysis purposes (*e.g.*, hydrogeological modelling).

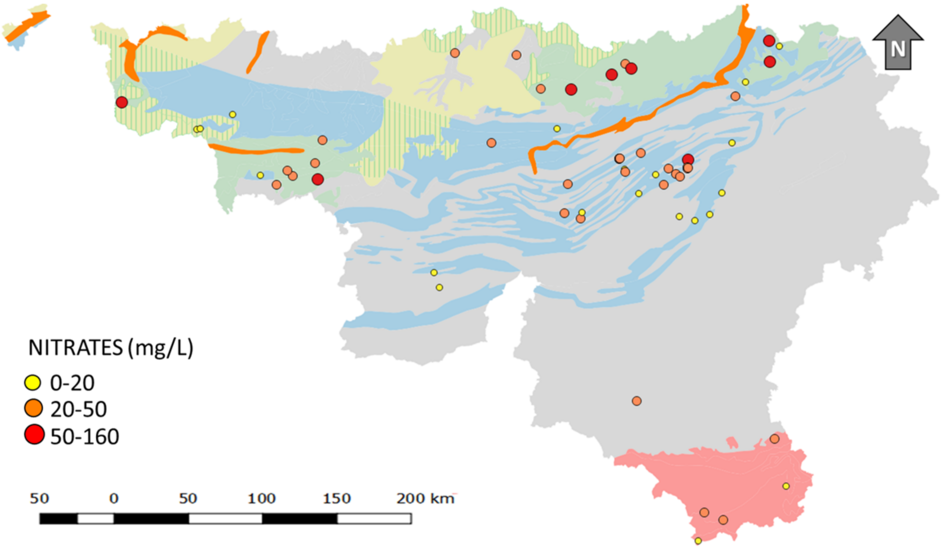


Fig. 5.10 Spatial distribution of the nitrate concentrations (mg/l) in groundwater for the spring 2016 campaign.

5.5 Conclusions

This paper presented the AkvaGIS GIS-based tool designed to improve the characterisation of groundwater bodies, with specific reference to analysing the availability and chemical quality of groundwater. The AkvaGIS tool was developed within the context of the FREEWAT project to include relevant information on groundwater quality and hydrogeological information in analysis of water resources.

The user-friendly and GIS-based architecture of AkvaGIS is significantly standardised and supplies an easy-to-use workflow that can manage, visualise and analyse hydrochemical and hydrogeological data. The AkvaGIS database structure ensures that all groundwater-related knowledge of a study area is archived and continuously updated without loss of the original information. Application of this tool can aid users in reinforcing the construction of conceptual models by cross-analysis of related data. In addition, AkvaGIS can simplify the preparation of input files for any groundwater numerical model in all of the available formats in QGIS.

An application of the AkvaGIS tools in the Walloon region (Belgium) demonstrated its usefulness by simplifying the steps needed to analyse the hydrochemical data. Use of analysis tools such as ionic balance analysis, Stiff maps, Piper diagrams, etc. facilitated understanding of the hydrochemical relationship among aquifers in the region and deduction of selected preliminary characteristics. This process represents a first step in further analysis of the region by the scientific community, public administration and the private sector for a wide range of environmental projects (*e.g.*, water supply, water quality control, mining control, among others). In addition, these first observations might spur future strategies focused on continued preservation of water quality and quantity indices in the Walloon region.

AkvaGIS aims to endorse water management and planning by simplifying the application of water-related directives (*e.g.*, Water Framework Directive) focusing on groundwater bodies. The scientific community, water resource authorities, and the private sector might benefit from using AkvaGIS, thus reducing the costs of commercial software and improving open sharing of hydrochemical and hydrogeological data and its interpretations in the water governance process.

Due to its open-source architecture, AkvaGIS can be updated and extended depending on the tailored applications. The FREEWAT community ensures proper functionality of all tools, manuals and their training material. Further development will address hydrochemical and hydrogeological analysis from different aspects such as a better connection between AkvaGIS and the hydrochemical numerical models.

Chapter 6

Conclusions and Outlook

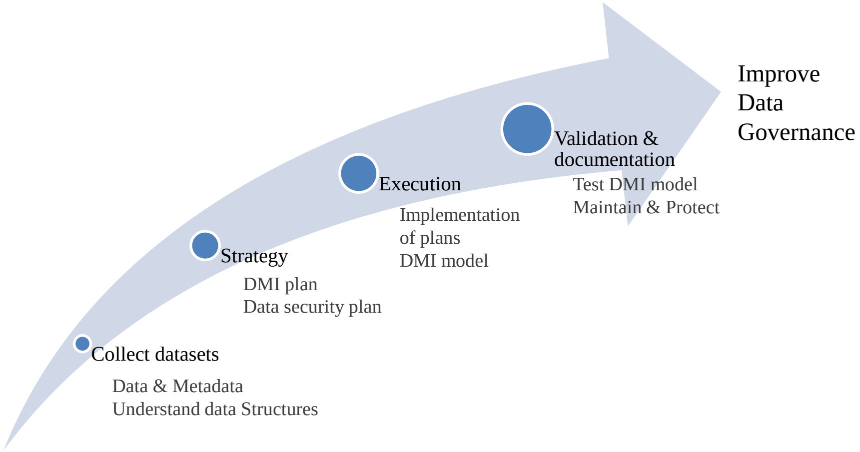
Conclusions and Outlook

This thesis has been developed in the framework of the Industrial Doctorates Plan ("Doctorat Industrial") in agreement among the Barcelona Water Cycle ("Barcelona Cicle de l'Aigua, S.A.", BCASA), from the Barcelona city Council; the Universitat Politècnica de Catalunya (UPC) and the Spanish Council for Scientific Research (IDAEA-CSIC).

The aim of this thesis has been designing, developing and applying methods and tools to improve management, migration, integration, analysis and interpretation of hydrogeological data.

Questions addressed along this dissertation and their answers are briefly summarised below:

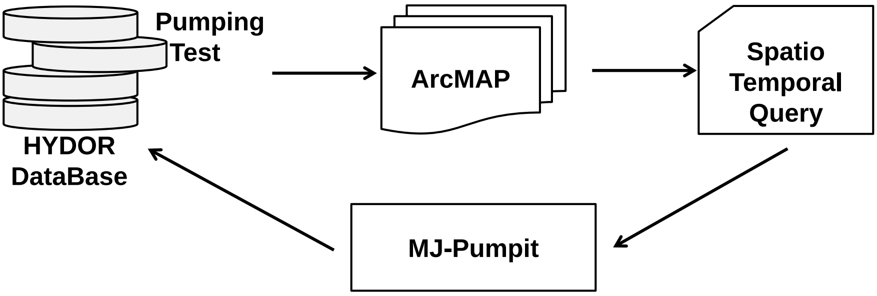
i Data infrastructures (for example, spatial data infrastructures) and their governance are continuously being used and developed because of their impact on groundwater management. A key factor in this trend is the need to reduce complexity in data migration and integration (DMI) projects and processes. This requeriments are adrresing in this thesis providing dynamic and scalable methodologies for migrating and integrating multiple infrastructures through a novel data migration and integration (DMI) methodology. This DMI methodology has been implemented in the Barcelona city Council increasing the volume of quality data to improve the understanding of the groundwater system behaviour and improves the monitoring network in the city. Other benefits observed during their implementation are that DMI methodology aid to check the data quality in original sources of information and to encourage collaboration among people with different skills from different departments. The methodology proposed can be widely implemented in any kind of DMI process, to develop Data warehouses, Spatial Data Infrastructures or to implement ICT tools on existing data infrastructures for further analyses improving data governance.



ii To gain higher performance of their analysis within their spatial context has been addressed by the implementation of the last methodology in Barcelona city Council systems to connect them with ICT GIS-based tools. The ICT tools implemented facilitate the manipulation, visualisation and analysis of the hydrogeological data in a unique GIS-based platform (HYDORGIS tools) while the original data infrastructures continue storing groundwater information. In addition, these set of GIS-based tools are being implemented as teaching subject and applied in widely environmental projects such as groundwater management, civil works, mining, among others.

iii To provide specific tools to analyse hydrogeological processes and to obtain hydraulic parameters have been performed by:

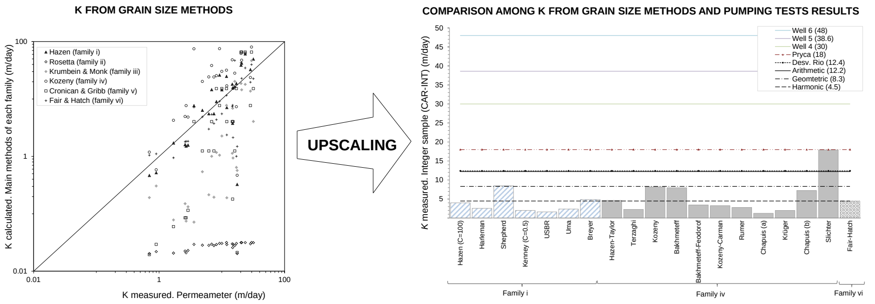
  a The design and development of a set of tools to integrate different instruments to collect, manage, process, represent and analyse data derived from pumping tests analyses, through: (1) an interface that improves interaction with a pumping test interpretation computer code in a platform that has extensive tools for the analysis and visualisation of data (MJ-Pumpit), (2) tools for advanced temporal-spatial analysis built in a GIS environment (HYYH Toolbar included in ArcMAP, ESRI) and (3) a geospatial database to maintain complete control of data involved in pumping tests.



This set of tools constitutes an operational and user-friendly analysis platform, which exploits several analysis and visualisation tools to ensure optimal results in hydrogeological studies. These GIS-based tools allow users to perform additional interpretations by considering different models, accurate geometries (*e.g.*, boundary conditions) of the study area and built-in plotting tools. The interconnection among the MJ-Pumpit, GIS environment and database ensures the proper management of a hydrogeological study. In addition, Appendix A shows more in detail registration of these tools and institutions where have been or are being used, to our knowledge.

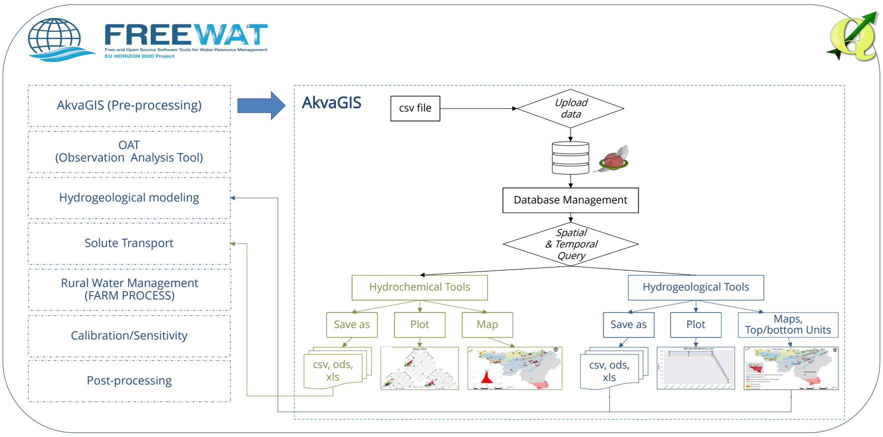
b The estimation of the hydraulic conductivity (K) through methods based on grain size at the centimetre scale and their further use in high scale studies. After the validation of each method by comparison between their results and the permeameter measurements of K, their application in high scales has been carried out by comparison among results of upscaling with pumping tests performed in the study area. In general, methods evaluated are highly correlated among them, but are not satisfactorily correlated with the permeameter measures of K. Most of the formulas give K values lower than measured K, so it leads to underestimation of this parameter. The comparison among methods evaluated using upscaling with pumping tests performed in the study area, shows the validation of the upscaling methodology to understand the behaviour of the flow regime.

The confidence of the Slichter method (1899) used to obtain an equivalent hydraulic conductivity in sandy samples to obtain an equivalent hydraulic conductivity, fits with results obtained with field tests (which have a greater scale of observation). These first results could be a step forward in the application of these equations at lager scale by applying them in an integrated sample. For instance, this methodology can be applied as a first step in the hydraulic parameterization of hydrogeological studies in an inexpensive manner.



iv To share open-source and user-friendly software designed to improve the characterisation of groundwater bodies, with specific reference to analysing the availability and chemical quality of groundwater. This new generation of GIS-based tool (AkvaGIS) was developed within the context of the FREEWAT H2020 project to include relevant information on groundwater quality and hydrogeological information in analysis of water resources. It allows to standardise, manage, analyse and interpret hydrogeological and hydrochemical data. In addition, using AkvaGIS can be used to prepare input files for most groundwater flow numerical models (*e.g.*, MODFLOW implemented into the FREEWAT platform) in all of the available formats in QGIS. The user-friendly and GIS-based architecture of AkvaGIS is significantly standardised and supplies an easy-to-use workflow that can manage, visualise and analyse hydrochemical and hydrogeological data, as their application in the Walloon Region shows. Due to its open-source architecture, AkvaGIS can be updated and extended depending on the tailored applications.

To further this research, the author propose to develop and maintain the GIS-based tools themselves (including data model) and design new tools to manage, visualise, analyse and interpret other environmental issues related with water cycle,



such as recharge of aquifers or further analyses by isotopes, always fed by the same spatial database. Therefore, using the DMI methodology proposed to connect all these set of GIS-based tools with external systems would be further enrichment for enhancing of hydrogeological studies. The scientific community, water resource authorities, and the private sector may be benefit of these connections without losing original sources of information and, thus, optimise the subsurface management.

More broadly, the author believe that this research will serve as a base for future enhancements on their using as training within living labs. Their results and users feedback would improve tools themselves and gives a better understanding on how these tools may facilitate communication among different technical skills and departments to optimise hydrogeological projects and reducing uncertainties in the decision process among them.

Regarding the specific tools to analyse hydrogeological processes and to obtain hydraulic parameters, it would be interesting further evaluation of the methods compiled and the upscaling methodology in other hydrogeological contexts.

References

AGAUR. Doctorats industrials, 2019. URL `http://doctoratsindustrials.gencat.cat/en/contents/view/6`.

M. Alcaraz, E. Vázquez-Suñé, V. Velasco, and R. Criollo. A loosely coupled GIS and hydrogeological modeling framework. *Environmental Earth Sciences*, 76(11): 382, 6 2017. ISSN 1866-6280. doi: 10.1007/s12665-017-6709-3. URL `http://link.springer.com/10.1007/s12665-017-6709-3`.

M. d. M. Alcaraz. *GIS platform for management of shallow geothermal resources.* PhD thesis, Universitat Politècnica de Catalunya (UPC), 3 2016. URL `https://tdx.cat/handle/10803/393738`.

E. E. Alonso, A. Gens, and A. Lloret. Precompression design for secondary settlement reduction. *Géotechnique*, 50(6):645–656, 12 2000. ISSN 0016-8505. doi: 10.1680/geot.2000.50.6.645. URL `http://www.icevirtuallibrary.com/doi/10.1680/geot.2000.50.6.645`.

P. Altheide. Spatial Data Transfer Standard (SDTS). In *Encyclopedia of GIS*, pages 1087–1095. Springer US, Boston, MA, 2008. doi: 10.1007/978-0-387-35973-1{_}1259. URL `http://link.springer.com/10.1007/978-0-387-35973-1_1259`.

M. S. Alyamani and Z. Sen. Determination of Hydraulic Conductivity from Complete Grain-Size Distribution Curves. *Ground Water*, 31(4):551–555, 1993.

K. M. Andreadis, N. Das, D. Stampoulis, A. Ines, J. B. Fisher, S. Granger, J. Kawata, E. Han, and A. Behrangi. The Regional Hydrologic Extremes Assessment System: A software framework for hydrologic modeling and data assimilation. *PLOS ONE*, 12(5):e0176506, 5 2017. ISSN 1932-6203. doi: 10.1371/journal.pone.0176506. URL `http://dx.plos.org/10.1371/journal.pone.0176506`.

S. Ashraf, H. Afshari, and A. G. Ebadi. American journal of agricultural and biological sciences. *American Journal of Agricultural and Biological Science*, 2011. URL `http://agris.fao.org/agris-search/search.do?recordID=DJ2012079623`.

I. S. Babiker, M. A. A. Mohamed, and T. Hiyama. Assessing groundwater quality using GIS. *Water Resources Management*, 21(4):699–715, 2 2007. ISSN 0920-4741. doi: 10.1007/s11269-006-9059-6. URL `http://link.springer.com/10.1007/s11269-006-9059-6`.

B. Bakhmeteff and N. Feodoroff. Flow through granular media. *Journal of Applied Mechanics*, 4:97–104, 1937.

References

M. Bakker. Wigaem An Analytic Element Model For Periodic Flow, 2009. URL `www.code.google.com/p/wigaem`.

P. Barlow and A. Moench. WTAQ—A computer program for calculating drawdowns and estimating hydraulic properties for confined and water-table aquifers. Technical report, Water-Resources Investigations. USGS U.S. Geological Survey, Northborough, Massachusetts, US, 1999. URL `http://pubs.usgs.gov/wri/wri99-4225/pdf/wri99-4225.pdf`.

J. Bear. *Dynamics of fluids in porous media.* Number 576. American Elsevier Publishing Company, New York. U S A, 1972.

J. Bear. *Hydraulics of Groundwater.* McGraw-Hill International Book Co., New York, 1979. ISBN 0-486-45355-3.

G. Bhatt, M. Kumar, and C. J. Duffy. Bridging the Gap between Geohydrologic Data and Distributed Hydrologic Modeling, 2008.

G. Bhatt, M. Kumar, and C. J. Duffy. A tightly coupled GIS and distributed hydrologic modeling framework. *Environmental Modelling & Software*, 62:70–84, 12 2014. ISSN 1364-8152. doi: 10.1016/J.ENVSOFT.2014.08.003. URL `https://www.sciencedirect.com/science/article/pii/S1364815214002266`.

D. L. Blodgett, N. L. Booth, T. C. Kunicki, J. I. Walker, and R. J. Viger. Description and testing of the Geo Data Portal: Data integration framework and Web processing services for environmental science collaboration. Technical report, USGS, 2011. URL `https://pubs.er.usgs.gov/publication/ofr20111157`.

S. J. Blott and K. Pye. Particle size distribution analysis of sand-sized particles by laser diffraction: an experimental investigation of instrument sensitivity and the effects of particle shape. *Sedimentology*, 53(3):671–685, 5 2006. doi: 10.1111/j.1365-3091.2006.00786.x. URL `http://doi.wiley.com/10.1111/j.1365-3091.2006.00786.x`.

E. Boisvert and B. Brodaric. GroundWater Markup Language (GWML) – enabling groundwater data interoperability in spatial data infrastructures. *Journal of Hydroinformatics*, 14(1):93 – 107, 1 2012. ISSN 1464-7141. doi: 10.2166/hydro.2011.172. URL `http://jh.iwaponline.com/cgi/doi/10.2166/hydro.2011.172`.

E. Boisvert, B. Brodaric, H. Julien, A. Smirnoff, and F. Létourneau. The GIN Mediator : a software tool for enabling interoperability in groundwater data networks. Technical Report Figure 1, Boisver and Brodaric, 2007.

L. A. Bojórquez-Tapia, S. Diaz-Mondragón, and E. Ezcurra. GIS-based approach for participatory decision making and land suitability assessment. *International Journal of Geographical Information Science*, 15(2):129–151, 3 2001. ISSN 1365-8816. doi: 10.1080/13658810010005534. URL `http://www.tandfonline.com/doi/abs/10.1080/13658810010005534`.

J. Bouma, C. Belmans, L. W. Dekker, and W. J. M. Jeurissen. Assessing the Suitability of Soils With Macropores for Subsurface Liquid Waste Disposal1. *Journal of Environment Quality*, 12(3):305, 1983. ISSN 0047-2425. doi: 10.2134/jeq1983.00472425001200030002x. URL `https://www.agronomy.org/publications/jeq/abstracts/12/3/JEQ0120030305`.

H. Bouwer and R. C. Rice. A Slug Test for Determining Hydraulic Conductivity of Unconfined Aquifers With Completely or Partially Penetrating Wells. *Water Resources Research*, 12(3):423 – 428, 6 1976. doi: 10.1029/WR012i003p00423.

Brahy V. Les indicateurs clés de l'environnement wallon 2014. Technical report, Service public de Wallonie., Namur, Belgique, 2014.

C. V. Brown. Examining the Emergence of Hybrid IS Governance Solutions: Evidence From a Single Case Site. *Information Systems Research*, 8(1):69–94, 3 1997. doi: 10.1287/isre.8.1.69. URL `http://pubsonline.informs.org/doi/abs/10.1287/isre.8.1.69`.

BSD. Berkeley Software Distribution License, 2017. URL `https://opensource.org/licenses/BSD-2-Clause`.

A. H. M. M. A. Cabalska, Jolanta; Felter. The Polish Hydrogeological Survey Database Integrator — a new GIS tool for the hydrogeological database management useful in mapping process, 2005. URL `http://www.pgi.gov.pl/images/stories/przeglad/pdf/pg_2005_10_2_10.pdf`.

E. Cabello, P. . Cuevas, and J. L. . Ramos. 3D modelling of grain size distribution in Quaternary deltaic deposits (Llobregat Delta, NE Spain). *Geologica Acta*, 5(3): 231–244, 2007. URL `http://www.redalyc.org/articulo.oa?id=50550302`.

J. Carbonell, A. Pérez Paricio, and J. Carrera. MariaJ_IV: programa de calibración automática de ensayos de bombeo: modelos analíticos y numéricos para medios 2-D y 3-D, 1997. URL `https://h2ogeo.upc.edu/es/investigacion-hidrologia-subterrania/software/34-ephebo`.

C. J. Carleton, R. a. Dahlgren, and K. W. Tate. A relational database for the monitoring and analysis of watershed hydrologic functions: I. Database design and pertinent queries. *Computers & Geosciences*, 31(4):393–402, 5 2005. ISSN 00983004. doi: 10.1016/j.cageo.2004.10.007. URL `http://linkinghub.elsevier.com/retrieve/pii/S0098300404002122`.

J. Carrera, E. Vázquez-Suñé, O. Castillo, and X. Sánchez-Vila. A methodology to compute mixing ratios with uncertain end-members. *Water Resources Research*, 40(12), 12 2004. ISSN 00431397. doi: 10.1029/2003WR002263. URL `http://doi.wiley.com/10.1029/2003WR002263`.

J. Carrera-Hernández and S. Gaskin. The Basin of Mexico Hydrogeological Database (BMHDB): Implementation, queries and interaction with open source software. *Environmental Modelling & Software*, 23(10-11):1271–1279, 10 2008. ISSN 13648152. doi: 10.1016/j.envsoft.2008.02.012. URL `http://linkinghub.elsevier.com/retrieve/pii/S1364815208000418`.

W. D. Carrier. Goodbye, Hazen; Hello, Kozeny-Carman. *Journal of Geotechnical and Geoenvironmental Engineering*, 129(11):1054–1056, 11 2003. ISSN 1090-0241. doi: 10.1061/(ASCE)1090-0241(2003)129:11(1054). URL `http://ascelibrary.org/doi/abs/10.1061/%28ASCE%291090-0241%282003%29129%3A11%281054%29`.

R. P. Chapuis. Predicting the saturated hydraulic conductivity of sand and gravel using effective diameter and void ratio. *Canadian Geotechnical Journal*, 41(5): 787–795, 2004. doi: 10.1139/T04-022.

R. P. Chapuis. Estimating the in situ porosity of sandy soils sampled in boreholes. *Engineering Geology*, 141-142:57–64, 7 2012. ISSN 00137952. doi: 10.1016/j.enggeo.2012.04.015. URL `http://linkinghub.elsevier.com/retrieve/pii/S0013795212001706http://www.sciencedirect.com/science/article/pii/S0013795212001706`.

R. P. Chapuis, F. Duhaime, and E. M. Benabdallah. Monitoring wells in clay: the apparently static water level and its influence during variable-head permeability tests. *Bulletin of Engineering Geology and the Environment*, 71(4):663–678, 9 2012. ISSN 1435-9529. doi: 10.1007/s10064-012-0433-8. URL `http://link.springer.com/10.1007/s10064-012-0433-8`.

D. Chen, S. Shams, C. Carmona-Moreno, and A. Leone. Assessment of open source GIS software for water resources management in developing countries. *Journal of Hydro-environment Research*, 4(3):253–264, 10 2010. ISSN 1570-6443. doi: 10.1016/J.JHER.2010.04.017. URL `https://www.sciencedirect.com/science/article/pii/S1570644310000511`.

I. Chenini and A. Ben Mammou. Groundwater recharge study in arid region: An approach using GIS techniques and numerical modeling. *Computers & Geosciences*, 36(6):801–817, 6 2010. doi: 10.1016/J.CAGEO.2009.06.014.

J.-Y. Cheong, S.-Y. Hamm, H.-S. Kim, E.-J. Ko, K. Yang, and J.-H. Lee. Estimating hydraulic conductivity using grain-size analyses, aquifer tests, and numerical modeling in a riverside alluvial system in South Korea. *Hydrogeology Journal*, 16 (6):1129–1143, 5 2008. ISSN 1431-2174. doi: 10.1007/s10040-008-0303-4. URL `http://apps.webofknowledge.com/full_record.do?product=UA&search_mode=GeneralSearch&qid=1&SID=P2H6Dggfkc4DN8BF1KH&page=1&doc=1`.

R. Chesnaux, M. Lambert, J. Walter, U. Fillastre, M. Hay, A. Rouleau, R. Daigneault, A. Moisan, and D. Germaneau. Building a geodatabase for mapping hydrogeological features and 3D modeling of groundwater systems: Application to the Saguenay–Lac-St.-Jean region, Canada. *Computers & Geosciences*, 37(11):1870–1882, 11 2011. doi: 10.1016/j.cageo.2011.04.013.

E. Commission. Council Directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption. Technical report, European Commission, Brussels, Belgium, 1998. URL `https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A31998L0083`.

E. Commission. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for community action in the field of water policy. Off. J. L, 327 22/12/2000. Technical report, European Commission, Brussels, Belgium, 2000.

R. Criollo, V. Velasco, E. Vázquez-Suñé, A. Serrano-Juan, M. Alcaraz, and A. García-Gil. An integrated GIS-based tool for aquifer test analysis. *Environmental Earth Sciences*, 75(5):391, 2 2016. ISSN 1866-6280. doi: 10.1007/s12665-016-5292-3. URL `http://link.springer.com/10.1007/s12665-016-5292-3`.

R. Criollo, V. Velasco, E. Vázquez-Suñé, A. Nardi, M. Marazuela, R. Rossetto, I. Borsi, L. Foglia, M. Cannata, and G. De Filippis. Open source GIS based tools to improve hydrochemical water resources management in EU H2020 FREEWAT platform. In *EGU General Assembly 2017*, volume 19, pages 164–180, Wien, Austria, 2017. URL `https://meetingorganizer.copernicus.org/EGU2017/EGU2017-19584.pdf`.

R. Criollo, E. Vázquez-Suñé, C. Riera, N. Kukuric, I. Borsi, L. Foglia, M. Cannata, G. De Filippis, and R. Rossetto. Water management capacity building through the FREEWAT platform. In *Geophysical Research Abstracts. EGU2018*, volume 20, Wien, Austria, 2018a. URL `https://meetingorganizer.copernicus.org/EGU2018/EGU2018-18956.pdf`.

R. Criollo, V. Velasco, S. Brouyère, C. Riera, E. Pujades, R. Rossetto, E. Vázquez-Suñé, and A. Jurado. AkvaGIS application in the Walloon region (Belgium). Technical report, OSF, 2018b. URL `https://osf.io/7urc9/`.

A. E. Cronican and M. M. Gribb. Hydraulic Conductivity Prediction for Sandy Soils. *Ground Water*, 42(3):459–64, 5 2004. ISSN 0017-467X. doi: 10.1111/j.1745-6584.2004.tb02694.x. URL `http://www.ncbi.nlm.nih.gov/pubmed/15161164http://doi.wiley.com/10.1111/j.1745-6584.2004.tb02694.x`.

E. Custodio and M. R. Llamas. *Hidrologia subterránea.* Omega, 1984.

G. Dagan. *Flow and Transport in Porous Formations.* Springer-Verlag Berlin Heidelberg, 1989. ISBN 978-3-642-75015-1. doi: 10.1007/978-3-642-75015-1.

J.-R. de Dreuzy, J. Bodin, H. Le Grand, P. Davy, D. Boulanger, A. Battais, O. Bour, P. Gouze, and G. Porel. General database for ground water site information. *Ground water*, 44(5):743–8, 2006. ISSN 0017-467X. doi: 10.1111/j.1745-6584.2006.00220.x. URL `http://www.ncbi.nlm.nih.gov/pubmed/16961496`.

J.-R. de Dreuzy, P. de Boiry, G. Pichot, and P. Davy. Use of power averaging for quantifying the influence of structure organization on permeability upscaling in on-lattice networks under mean parallel flow. *Water Resources Research*, 46 (8):W08519, 8 2010. ISSN 00431397. doi: 10.1029/2009WR008769. URL `http://doi.wiley.com/10.1029/2009WR008769`.

G. De Filippis, I. Borsi, L. Foglia, M. Cannata, V. Velasco Mansilla, E. Vasquez-Suñe, M. Matteo Ghetta, and R. Rossetto. Software tools for sustainable water resources management: the GIS-integrated FREEWAT platform. *Rendiconti Online della Società Geologica Italiana*, 42:59–61, 3 2017. ISSN 20358008. doi: 10.3301/ROL.2017.14. URL `http://rendiconti.socgeol.it/244/fulltext.html?ida=3853`.

G. De Filippis, Borsi I., Foglia L., M. Cannata, R. Criollo, E. Vázquez-Suñé, I. Kopač, B. Panteleit, P. Positano, M. Nannucci, M. Sapiano, D. Svidzinska, M. Grodzynskyi, and R. Rossetto. Joining participatory approach and spatially-based modelling tools for groundwater resource management. In *Geophysical Research Abstracts. EGU2018*, volume 20, Wien, Austria, 2018. URL `https://meetingorganizer.copernicus.org/EGU2018/EGU2018-8396.pdf`.

Y. T. Dile, P. Daggupati, C. George, R. Srinivasan, and J. Arnold. Introducing a new open source GIS user interface for the SWAT model. *Environmental Modelling and Software*, 85:129–138, 11 2016. ISSN 13648152. doi: 10.1016/j.envsoft.2016.08.004. URL `https://www.sciencedirect.com/science/article/pii/S1364815216304674http://dx.doi.org/10.1016/j.envsoft.2016.08.004`.

Diputación de Alicante. PIBE 2.0. User manual, 2006. URL `https://ciclohidrico.com/descargas/aplicaciones/https://docs.google.com/file/d/0B1MJStxDOT8ba1YxMjZjY1NDWXM/edit`.

References

L. Duarte and J. Gonçalves. Open source GIS tools: two environmental applications. In A. Teodoro, editor, *Frontiers in Information Systems. GIS – An Overview of Applications*, chapter 4. Bentham Science Publishers, 2018. ISBN 978-1-68108-612-5.

G. M. Duffield. AqTestSolv |Aquifer Test. Summary of Aquifer Test Solutions, 1989. URL `http://www.aqtesolv.com/summary_of_solutions.htmhttp://www.slb.com/services/additional/water/software/pumping_test.aspx`.

Y. Erzin, S. D. Gumaste, A. K. Gupta, and D. N. Singh. Artificial neural network (ANN) models for determining hydraulic conductivity of compacted fine-grained soils. *Canadian Geotechnical Journal*, 46(8):955–968, 8 2009. ISSN 0008-3674. doi: 10.1139/T09-035. URL `http://www.nrcresearchpress.com/doi/abs/10.1139/t09-035#.USn6YDCQVZ4`.

G. Eshel, G. J. Levy, U. Mingelgrin, and M. J. Singer. Critical Evaluation of the Use of Laser Diffraction for Particle-Size Distribution Analysis. *Soil Science Society of America Journal*, 68(3):736–743, 2004. ISSN 1435-0661. doi: 10.2136/sssaj2004.7360. URL `https://www.crops.org/publications/sssaj/abstracts/68/3/736`.

ESRI. ArcGIS desktop developer guide: ArcGIS 9.1, 2005.

ESRI. ArcGIS 10. Environmental Systems Research Institute. Technical report, Environmental Systems Research Institute (ESRI), Redlands, United States of America, 2012. URL `http://desktop.arcgis.com/en/`.

ESRI. Data migration from one storage type to another—Help | ArcGIS Desktop, 2018. URL `http://desktop.arcgis.com/en/arcmap/10.3/manage-data/gdbs-in-oracle/data-migration.htm`.

G. M. Fair, L. P. Hatch, E. Herbert, and J. Hudson. *Fundamental factors governing the streamline flow of water through sand*, volume 25. American Water Works Association, 1933. doi: 10.2307/41225921.

P. Fitch, B. Brodaric, M. Stenson, N. Booth, P. Fitch, and b. M. Stenson. Integrated Groundwater Data Management. *Integrated Groundwater Management*, pages 667 – 692, 2016. doi: 10.1007/978-3-319-23576-9{_}26. URL `http://link.springer.com/10.1007/978-3-319-23576-9_26https://link.springer.com/content/pdf/10.1007/978-3-319-23576-9_26.pdf`.

L. Foglia, I. Borsi, M. Cannata, G. De Filippis, R. Criollo, S. Mehl, and R. Rossetto. How innovative ICT tools can enhance understanding of interactions between societal, hydrological and environmental changes. In *AGU 2017 Fall Meeting*, New Orleans, USA, 2017. URL `https://agu.confex.com/agu/fm17/meetingapp.cgi/Paper/272696`.

L. Foglia, I. Borsi, S. Mehl, G. De Filippis, M. Cannata, E. Vasquez-Suñe, R. Criollo, and R. Rossetto. FREEWAT, a Free and Open Source, GIS-Integrated, Hydrological Modeling Platform. *Groundwater*, 56(4):521–523, 2018. ISSN 17456584. doi: 10.1111/gwat.12654.

D. Gámez, J. A. Simó, F. J. Lobo, A. Barnolas, J. Carrera, and E. Vázquez-Suñé. Onshore-offshore correlation of the Llobregat deltaic system, Spain: Development of deltaic geometries under different relative sea-level and growth fault influences. *Sedimentary Geology*, 217(1-4):65–84, 2009. ISSN 00370738. doi: 10.1016/j.sedgeo.2009.03.007. URL `http://dx.doi.org/10.1016/j.sedgeo.2009.03.007`.

GE Solutions. Smallworld GIS software, 2018. URL `https://www.gegridsolutions.com/geospatial/catalog/smallworld_core.htm`.

L. W. Gelhar, C. Welty, and K. R. Rehfeldt. A critical review of data on field-scale dispersion in aquifers. *Water Resources Research*, 28(7):1955–1974, 7 1992. ISSN 00431397. doi: 10.1029/92WR00607. URL `http://doi.wiley.com/10.1029/92WR00607`.

L. Gelhar W. and C. Axness L. Three-dimensional stochastic analysis of macrodispersion in aquifers. *Water Resources Research*, 19(1):161–180, 1983.

GeoSoft. Well32 Software developed by geo&soft, 2014. URL `https://www.geoandsoft.com/`.

A. Ghosh, A. K. Tiwari, and S. Das. A GIS based DRASTIC model for assessing groundwater vulnerability of Katri Watershed, Dhanbad, India. *Modeling Earth Systems and Environment*, 1(3):11, 10 2015. ISSN 2363-6203. doi: 10.1007/s40808-015-0009-2. URL `http://link.springer.com/10.1007/s40808-015-0009-2`.

T. Gleeson, W. M. Alley, D. M. Allen, M. A. Sophocleous, Y. Zhou, M. Taniguchi, and J. VanderSteen. Towards Sustainable Groundwater Use: Setting Long-Term Goals, Backcasting, and Managing Adaptively. *Ground Water*, 50(1):19–26, 1 2012. ISSN 0017467X. doi: 10.1111/j.1745-6584.2011.00825.x. URL `http://doi.wiley.com/10.1111/j.1745-6584.2011.00825.x`.

T. Gleeson, K. M. Befus, S. Jasechko, E. Luijendijk, and M. B. Cardenas. The global volume and distribution of modern groundwater. *Nature Geoscience*, 9(2):161–164, 2016a. ISSN 17520908. doi: 10.1038/ngeo2590.

T. Gleeson, K. M. Befus, S. Jasechko, E. Luijendijk, and M. B. Cardenas. The global volume and distribution of modern groundwater. *Nature Geoscience*, 9 (2):161–167, 2 2016b. ISSN 1752-0894. doi: 10.1038/ngeo2590. URL `http://www.nature.com/articles/ngeo2590`.

GNU. Lesser General Public License v2.0., 2017. URL `https://www.gnu.org/licenses/old-licenses/lgpl-2.0.html`.

R. Gogu, G. Carabin, V. Hallet, V. Peters, and A. Dassargues. GIS-based hydrogeological databases and groundwater modelling. *Hydrogeology Journal*, 9(6):555–569, 12 2001. ISSN 1431-2174. doi: 10.1007/s10040-001-0167-3. URL `http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s10040-001-0167-3`.

M. F. Goodchild. Twenty years of progress: GIScience in 2010. *Journal of Spatial Information Science*, 1(1):3–20, 2010. ISSN 1948-660X. doi: 10.5311/JOSIS.2010.1.2. URL `http://josis.org/index.php/josis/article/view/32`.

M. F. Goodchild. 1.03 - Big Geodata. In *Comprehensive Geographic Information Systems*, pages 19–25. Elsevier, 1 2018. ISBN 978-0-12-804793-4. doi: 10.1016/B978-0-12-409548-9.09595-6.

S. C. Guptill, J. L. Morrison, and International Cartographic Association. *Elements of spatial data quality.* Elsevier Science, 1995. ISBN 9781483287942.

K. Halford and E. Kunianksy. USGS OFR 02-197 Spreadsheets for the Analysis of Aquifer Pumping and Slug Test Data. Technical report, USGS U.S. Geological Survey, 2002. URL `http://pubs.usgs.gov/of/2002/ofr02197/`.

A. Hamblin. Sustainable agricultural systems - What are the appropriate measures for soil structure. *Australian Journal of Soil Research*, 29(6):709, 1991. doi: 10.1071/SR9910709.

A. Harbaugh. MODFLOW-2005, the U.S. Geological Survey Modular Groundwater Model - the Ground-water Flow Process. . U.S. Geological Survey, Techniques and Methods 6 -A16. Technical report, USGS, 2005.

R. A. Hazen. Some physical properties of sand and gravels, with special reference to their use in filtration: 24th Annual Rep., Massachusetts State Board of Health, 1 1892. URL `https://www.scienceopen.com/document?vid=199fad18-0c9d-4830-b920-cfbf680327c1`.

K. Hemker and V. Post. MLU for Windows Well flow modeling in multilayer aquifer systems UNESCO edition Aquifer and Pump Test Analysis Software MLU for Windows, 2010.

P. H. Hendriks, E. Dessers, and G. van Hootegem. Reconsidering the definition of a spatial data infrastructure. *International Journal of Geographical Information Science*, 26(8):1479–1494, 8 2012. ISSN 1365-8816. doi: 10.1080/13658816.2011.639301. URL `http://www.tandfonline.com/doi/abs/10.1080/13658816.2011.639301`.

Hitachi Vantara Community. Data Integration - Kettle | Hitachi Vantara Community, 2018. URL `https://community.hitachivantara.com/docs/DOC-1009855-data-integration-kettle`.

J. S. Horsburgh, D. G. Tarboton, M. Piasecki, D. R. Maidment, I. Zaslavsky, D. Valentine, and T. Whitenack. An integrated system for publishing environmental observations data. *Environmental Modelling & Software*, 24(8):879–888, 8 2009. ISSN 13648152. doi: 10.1016/j.envsoft.2009.01.002. URL `http://linkinghub.elsevier.com/retrieve/pii/S1364815209000036`.

P. Howard. Data Migration: A White Paper by Bloor Research. Technical report, Bloor research, 2008. URL `https://es.slideshare.net/findwhitepapers/data-migration-a-white-paper-by-bloor-research`.

B. Huang, T. Cova, M.-H. Tsou, G. Bareth, C. Song, Y. Song, K. Cao, and E. Silva. *Comprehensive Geographic Information Systems.* Elsevier, huang, bo edition, 2018. ISBN 9780128046609. URL `https://www.sciencedirect.com/referencework/9780128047934/comprehensive-geographic-information-systems`.

IGME. The chemical composition of natural groundwater (“La composición química de las aguas subterráneas naturales”). Technical report, IGME, 2002. URL `http://aguas.igme.es/igme/publica/libro43/pdf/lib43/1_1.pdf`.

INSPIRE. Infrastructure for Spatial Information in Europe. D2.9_V1.0. Guidelines for the use of Observations&Measurements and Sensor Web Enablements-related standards in INSPIRE. AnnexII and III data specification development. Technical Guidelines. Technical report, EU, 2011. URL `http://inspire.jrc.ec.europa.eu/`.

INSPIRE. Infrastructure for Spatial Information in Europe. D.2.8.11.4. Data specification on Geology-Draft. Technical Guidelines. Technical report, EU, 2013.

IRMA. *Geographic information systems: concepts, methodologies, tools, and applications.* Information Resources Management Association. Information Science Reference, 2013. ISBN 9781466620384.

ISO. ISO/IEC 38500:2008 - Corporate governance of information technology, 2008. URL `https://www.iso.org/standard/51639.html`.

O. H. Jarar, R. Löwner, L. Benaabidate, and J. Wächter. HydrIS: An open source GIS decision support system for groundwater management (Morocco). *Geo-Spatial Information Science*, 12(3):212–216, 1 2009. ISSN 10095020. doi: 10.1007/s11806-009-0048-9. URL `http://www.tandfonline.com/doi/abs/10.1007/s11806-009-0048-9`.

G. S. Johnson and D. M. Cosgrove. RADFLOW: A numerical model for pumping test analysis, 2001. URL `http://www.if.uidaho.edu/~johnson/usersmanual2.pdf`.

N. Jones, J. Nelson, G. Wiliams, and R. Hotchkiss. Groundwater Modeling System, 2004. URL `https://emrl.byu.edu/`.

A. Journel, C. Deutsch, and A. Desbarats. Power Averaging for Block Effective Permeability. In *SPE California Regional Meeting.* Society of Petroleum Engineers, 4 1986. ISBN 978-1-55563-615-9. doi: 10.2118/15128-MS. URL `http://www.onepetro.org/doi/10.2118/15128-MS`.

J. Joy. Managing Data Separation and Migration During a Divestiture. In Management for Professionals, editor, *Divestitures and Spin-Offs*, pages 345–360. Springer, Boston, MA, 2018. doi: 10.1007/978-1-4939-7662-1{_}28. URL `http://link.springer.com/10.1007/978-1-4939-7662-1_28`.

A. Jurado, E. Vàzquez-Suñé, A. Soler, I. Tubau, J. Carrera, E. Pujades, and I. Anson. Application of multi-isotope data (O, D, C and S) to quantify redox processes in urban groundwater. *Applied Geochemistry*, 34:114–125, 7 2013. ISSN 08832927. doi: 10.1016/j.apgeochem.2013.02.018. URL `http://www.sciencedirect.com/science/article/pii/S0883292713000516`.

A. Jurado, P. Gago-Ferrero, E. Vàzquez-Suñé, J. Carrera, E. Pujades, M. S. Díaz-Cruz, and D. Barceló. Urban groundwater contamination by residues of UV filters. *Journal of hazardous materials*, 271:141–9, 4 2014. ISSN 1873-3336. doi: 10.1016/j.jhazmat.2014.01.036. URL `http://www.sciencedirect.com/science/article/pii/S0304389414000594`.

A. Jurado, E. Vázquez-Suñé, and E. Pujades. Potential uses of pumped urban groundwater: a case study in Sant Adrià del Besòs (Spain). *Hydrogeology Journal*, pages 1–14, 3 2017. ISSN 1431-2174. doi: 10.1007/s10040-017-1575-3. URL `http://link.springer.com/10.1007/s10040-017-1575-3`.

A. Jurado, A. V. Borges, E. Pujades, V. Hakoun, J. Otten, K. Knöller, and S. Brouyère. Occurrence of greenhouse gases in the aquifers of the Walloon Region (Belgium). *Science of The Total Environment*, 619-620:1579–1588, 4 2018. ISSN 00489697. doi: 10.1016/j.scitotenv.2017.10.144. URL `http://www.ncbi.nlm.nih.gov/pubmed/29128121https://linkinghub.elsevier.com/retrieve/pii/S0048969717328528`.

S. Kao, K. Ranatunga, G. Squire, A. Pratt, and D. Dee. Visualisation of hydrological observations in the water data transfer format. *Environmental Modelling & Software*, 26(12):1767–1769, 12 2011. ISSN 1364-8152. doi: 10.1016/J.ENVSOFT.2011.07.005. URL `https://www.sciencedirect.com/science/article/pii/S1364815211001666`.

M. Kasenow. *Determination of hydraulic conductivity from grain size analysis.* Water Resouces Publications, 2002. ISBN 9781887201315.

T. C. Kenney, D. Lau, and G. I. Ofoegbu. Permeability of compacted granular materials. *Canadian Geotechnical Journal*, 21(4):726–729, 11 1984. ISSN 0008-3674. doi: 10.1139/t84-080. URL `http://www.nrcresearchpress.com/doi/10.1139/t84-080`.

S.-M. Kim, Y. Choi, J. Suh, S. Oh, H.-D. Park, S.-H. Yoon, and W.-R. Go. ArcMine: A GIS extension to support mine reclamation planning. *Computers & Geosciences*, 46:84–95, 9 2012. ISSN 00983004. doi: 10.1016/j.cageo.2012.04.007. URL `http://www.sciencedirect.com/science/article/pii/S0098300412001306`.

A. Kingdon, M. L. Nayembil, A. E. Richardson, and A. G. Smith. A geodata warehouse: Using denormalisation techniques as a tool for delivering spatially enabled integrated geological information to geologists. *Computers & Geosciences*, 96:87–97, 2016. ISSN 00983004. doi: 10.1016/j.cageo.2016.07.016.

M. Konert and J. Vandenberghe. Comparison of laser grain size analysis with pipette and sieve analysis: a solution for the underestimation of the clay fraction. *Sedimentology*, 44(3):523–535, 6 1997. doi: 10.1046/j.1365-3091.1997.d01-38.x. URL `http://doi.wiley.com/10.1046/j.1365-3091.1997.d01-38.x`.

J. J. Korhonen, I. Melleri, K. Hiekkanen, and M. Helenius. Designing Data Governance Structure: An Organizational Perspective. *GSTF Journal on Computing*, 4(2), 2013. doi: 10.5176/2251-3043{_}2.4.203.

W. Korres and K. Schneider. 2.04 - GIS for Hydrology. In *Comprehensive Geographic Information Systems*, pages 51–80. Elsevier, 1 2018. ISBN 9780128047934. doi: 10.1016/B978-0-12-409548-9.09635-4. URL `http://linkinghub.elsevier.com/retrieve/pii/B9780124095489096354https://www.sciencedirect.com/science/article/pii/B9780124095489096354`.

J. Kozeny. *Hydraulics.* Elsevier, Amsterdam, 1953.

N. Kresic and A. Mikszewski. *Hydrogeological Conceptual Site Models. Data Analysis and Visualization.* CRC Press, 7 2012. ISBN 978-1-4398-5222-4. doi: 10.1201/b12151. URL `https://www.taylorfrancis.com/books/9781439852286`.

W. Krumbein and G. Monk. Permeability as a Function of the Size Parameters of Unconsolidated Sand. *Transactions of the AIME*, 151(01):153–163, 12 1943. doi: 10.2118/943153-G.

C. Lee, K. Kim, and H. Lee. GIS based optimal impervious surface map generation using various spatial data for urban nonpoint source management. *Journal of Environmental Management*, 206:587–601, 1 2018. ISSN 0301-4797. doi: 10.1016/J.JENVMAN.2017.10.076. URL `https://www.sciencedirect.com/science/article/pii/S0301479717310721`.

F. Leite, A. de Sousa, C. de Souza Baptista, C. Nunes, E. da Silva, D. de Almeida, and A. de Paiva. Migratool: Towards a Web-Based Spatial Database Migration Tool. In *16th International Workshop on Database and Expert Systems Applications (DEXA'05)*, pages 480–484. IEEE, 2005. ISBN 0-7695-2424-9. doi: 10.1109/DEXA.2005.132. URL `http://ieeexplore.ieee.org/document/1508319/`.

Létourneau F., Boisvert É., and Brodaric B. Groundwater Markup Language: A GML Application for the Exchange of Groundwater Data. *Geohydro 2011.ca. LHA, 2013*, 2011.

Z. Li, J. Quan, X.-Y. Li, X.-C. Wu, H.-W. Wu, Y.-T. Li, and G.-Y. Li. Establishing a model of conjunctive regulation of surface water and groundwater in the arid regions. *Agricultural Water Management*, 174:30–38, 8 2016. ISSN 0378-3774. doi: 10.1016/J.AGWAT.2016.04.030. URL `https://www.sciencedirect.com/science/article/pii/S0378377416301561`.

X. Luo, Y. Xu, and F. Zhou. Research on the integration of data warehouse, virtual reality and geographical information system in water resources management. In *Proceedings 2011 IEEE International Conference on Spatial Data Mining and Geographical Knowledge Services*, pages 497–500. IEEE, 6 2011. ISBN 978-1-4244-8352-5. doi: 10.1109/ICSDM.2011.5969095. URL `http://ieeexplore.ieee.org/document/5969095/`.

B. D. R. Maidment. CUAHSI Hydrologic Information Systems Symposium. In *A Data Model for Hydrologic Observations*, pages 1–23, Austin, Texas, 2005.

D. R. Maidment. Arc Hydro. GIS for Water Resources, 2002. URL `http://books.google.com/books?hl=es&lr=&id=07vH7Sf0v6MC&pgis=1http://www.crwr.utexas.edu/gis/archydrobook/ArcHydro.htm`.

A. Marchant, V. Banks, K. Royse, and S. Quigley. The development of a GIS methodology to assess the potential for water resource contamination due to new development in the 2012 Olympic Park site, London. *Computers & Geosciences*, 51: 206–215, 2 2013. ISSN 0098-3004. doi: 10.1016/J.CAGEO.2012.09.006. URL `https://www.sciencedirect.com/science/article/pii/S0098300412003159`.

T. J. Marshall. A relation between permeability and size distribution of pores. *Journal of Soil Science*, 9(1):1–8, 3 1958. doi: 10.1111/j.1365-2389.1958.tb01892.x. URL `http://doi.wiley.com/10.1111/j.1365-2389.1958.tb01892.x`.

A. Martens, M. Book, and V. Gruhn. A data decomposition method for stepwise migration of complex legacy data. In *Proceedings of the 40th International Conference on Software Engineering Software Engineering in Practice - ICSE-SEIP '18*, pages 33–42, New York, New York, USA, 2018. ACM Press. ISBN 9781450356596. doi: 10.1145/3183519.3183520. URL `http://dl.acm.org/citation.cfm?doid=3183519.3183520`.

Matplotlib 1.5. Python plotting — Matplotlib 1.5 documentation, 2017. URL `https://matplotlib.org/`.

K. Matthes, H. Nusche, P. Dietrich, and T. Vienken. Effects ofmeasuring inaccuracy during grain size analyses on the determination of hydraulic conductivity. *Grundwasser*, 17(2):105–111, 1 2012. ISSN 1430-483X. doi: 10.1007/s00767-012-0191-3. URL `http://www.springerlink.com/index/10.1007/s00767-012-0191-3`.

M. Mazzei and S. Di Guida. Spatial Data Warehouse and Spatial OLAP in Indoor/Outdoor Cultural Environments. In Gervasi O. et al., editor, *Computational Science and Its Applications – ICCSA 2018. ICCSA 2018. Lecture Notes in Computer Science*, pages 233–250. Springer, Cham, 5 2018. doi: 10.1007/978-3-319-95168-3{_}16. URL `http://link.springer.com/10.1007/978-3-319-95168-3_16`.

H. Merdun. Alternative methods in the development of pedotransfer functions for soil hydraulic characteristics. *Eurasian Soil Science*, 43(1):62–71, 2 2010. ISSN 1064-2293. doi: 10.1134/S1064229310010084. URL `http://www.springerlink.com/index/10.1134/S1064229310010084`.

V. Merwade, A. Cook, and J. Coonrod. GIS techniques for creating river terrain models for hydrodynamic modeling and flood inundation mapping. *Environmental Modelling & Software*, 23(10-11):1300–1311, 10 2008. ISSN 1364-8152. doi: 10.1016/J.ENVSOFT.2008.03.005. URL `https://www.sciencedirect.com/science/article/pii/S1364815208000455`.

Microsoft. Microsoft Office | Productivity Tools for Home & Office, 2010. URL `https://products.office.com/en-GB/?omkt=en-GB`.

Microsoft. Microsoft Developer Network online documentation, 2018. URL `https://msdn.microsoft.com/es-es/dn308572.aspx`.

MIT. Massachusetts Institute of Technology License, 2017. URL `https://opensource.org/licenses/MIT`.

C. Molano. Groundwater Spreadsheets: Efficient and Practical Resource for Solving Simple and Complex Flow, Pollution, and Environmental Problems. In NGWA, editor, *NGWA Event and Education Recordings Archive*. NGWA, 2013.

B. Nas and A. Berktay. Groundwater quality mapping in urban groundwater using GIS. *Environmental monitoring and assessment*, 160(1-4):215–27, 1 2010. ISSN 1573-2959. doi: 10.1007/s10661-008-0689-4. URL `http://www.ncbi.nlm.nih.gov/pubmed/19096909`.

A. Nielsen, K. Bolding, F. Hu, and D. Trolle. An open source QGIS-based workflow for model application and experimentation with aquatic ecosystems. *Environmental Modelling & Software*, 95:358–364, 9 2017. ISSN 13648152. doi: 10.1016/j.envsoft.2017.06.032. URL `http://linkinghub.elsevier.com/retrieve/pii/S1364815217301172`.

Odfpy 1.3. Odfpy PyPi 1.3. documentation, 2017. URL `https://pypi.python.org/pypi/odfpy`.

OGC. Observations and Measurements, 03-022r3. Technical report, OGC, 2003. URL `http://portal.opengeospatial.org`.

OGC. Observations and Measurements, 05-087r4. Technical report, OGC, 2006. URL `http://portal.opengeospatial.org`.

OGC. Observations and Measurements, Part 1, Observation schema, 07-022r1. Technical report, OGC, 2007. URL `http://www.opengeospatial.org/standards/om`.

OGC. Interoperability Experiment Report OGC. Groundwater Interoperability Experiment Final Report, 10-194r3. Technical report, OGC, 2011. URL `http://external.opengis.org/twiki_public/pub/HydrologyDWG/GroundwaterInteroperabilityExperiment/OGC_GWIE_Final_Report_Mar2011.pdf`.

OGC. WaterML 2.0:Part 1-Timeseries.10-126r3. Technical report, OGC, 2012. URL `https://portal.opengeospatial.org/files/?artifact_id=48531`.

OneGeology. OneGeology project, 2013. URL `http://www.onegeology.org/`.

Openpyxl 2.3. Documentation, 2017.

Oracle. Successful Data Migration. Technical report, Oracle, 2011. URL `https://www.oracle.com/technetwork/middleware/oedq/successful-data-migration-wp-1555708.pdf`.

B. Otto. A morphology of the organisation of data governance. Technical report, Association for Information Systems AIS Electronic Library (AISeL), 2011. URL `http://aisel.aisnet.org/ecis2011/272`.

H. J. Oulidi and A. Moumen. Towards a Spatial Data Infrastructure and an Integrated Management of Groundwater Resources. *Journal of Geographic Information System*, 07(06):667–676, 11 2015. ISSN 2151-1950. doi: 10.4236/jgis.2015.76054. URL `http://www.scirp.org/journal/doi.aspx?DOI=10.4236/jgis.2015.76054`.

C. C. Pasquale, B. Ionescu, D. Ionescu, V. diLecce, and A. Gurerero. Virtual Data Warehouse architecture for real-time WebGIS. In *2008 IEEE Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems*, pages 80–85. IEEE, 7 2008. ISBN 978-1-4244-1927-2. doi: 10.1109/VECIMS.2008.4592757. URL `http://ieeexplore.ieee.org/document/4592757/`.

A. Phillips, I. Williamson, C. Ezigbalike, and L. Williamson. Australian Surveyor Spatial Data Infrastructure Concepts Spatial Data Infrastructure Concepts. *The Auatralian Surveyor*, 44(1):20–28, 1999. doi: 10.1080/00050351.1999.10558768. URL `http://www.tandfonline.com/action/journalInformation?journalCode=tjss20`.

Pyexcel 0.2. Documentation, 2017. URL `https://pyexcel.readthedocs.io`.

PyQt4. PyQt4 PyPi Documentation, 2017. URL `https://pypi.org/project/PyQt4/`.

QGIS Development Team. QGIS Geographic Information System. Open Source Geospatial Foundation Project, 2009. URL `http://qgis.osgeo.org`.

K. Ranatunga, G. Walker, and P. A. Sheahan. Implementation of groundwater observations in the Water Data Transfer Format. *Environmental Modelling & Software*, 26(4):549–550, 4 2011. ISSN 1364-8152. doi: 10.1016/J.ENVSOFT.2010.10.005. URL `https://www.sciencedirect.com/science/article/pii/S1364815210002641`.

W. J. Rawls and D. L. Brakensiek. *Unsaturated Flow in Hydrologic Modeling*. Springer Netherlands, Dordrecht, 1989. ISBN 978-94-010-7559-6. doi: 10.1007/978-94-009-2352-2. URL `http://www.springerlink.com/index/10.1007/978-94-009-2352-2`.

J. C. Refsgaard, A. L. Højberg, I. Møller, M. Hansen, and V. Søndergaard. Groundwater modeling in integrated water resources management–visions for 2020. *Ground water*, 48(5):633–48, 2010. ISSN 1745-6584. doi: 10.1111/j.1745-6584.2009.00634.x. URL `http://www.ncbi.nlm.nih.gov/pubmed/19788560`.

X. Ren, Y. Zhao, Q. Deng, D. Li, and D. Wang. A relation of hydraulic conductivity – Void ratio for soils based on Kozeny-carman equation. *Engineering Geology*, 2016. ISSN 00137952. doi: 10.1016/j.enggeo.2016.08.017.

P. Renard. Hytool Installation Guide, 2008.

P. Renard and G. de Marsily. Calculating equivalent permeability: a review. *Advances in Water Resources*, 20(5-6):253–278, 1997. ISSN 03091708. doi: 10.1016/S0309-1708(96)00050-4. URL `http://linkinghub.elsevier.com/retrieve/pii/S0309170896000504`.

J. F. Rios, M. Ye, and L. Wang. uWATER-PA: Ubiquitous WebGIS Analysis Toolkit for Extensive Resources-Pumping Assessment. *Ground Water*, 49(6):776–780, 11 2011. ISSN 0017467X. doi: 10.1111/j.1745-6584.2011.00872.x. URL `http://doi.wiley.com/10.1111/j.1745-6584.2011.00872.x`.

J. F. Rios, M. Ye, L. Wang, P. Z. Lee, H. Davis, and R. Hicks. ArcNLET: A GIS-based software to simulate groundwater nitrate load from septic systems to surface water bodies. *Computers and Geosciences*, 52:1080–116, 3 2013. ISSN 00983004. doi: 10.1016/j.cageo.2012.10.003. URL `https://www.sciencedirect.com/science/article/pii/S0098300412003421`.

B. Rogiers, D. Mallants, O. Batelaan, M. Gedeon, M. Huysmans, and A. Dassargues. Estimation of Hydraulic Conductivity and Its Uncertainty from Grain-Size Data Using GLUE and Artificial Neural Networks. *Mathematical Geosciences*, 44 (6):739–763, 6 2012. ISSN 1874-8961. doi: 10.1007/s11004-012-9409-2. URL `http://www.springerlink.com/index/10.1007/s11004-012-9409-2`.

T. Rohrich. AquiferTest Pro | Pumping Test & Slug Test Analysis Software, 2002. URL `http://www.swstechnology.com/groundwater-software/pumping-test-analysis/aquifertest-prohttp://www.swstechnology.com/groundwater-software/pumping-test-analysis/aquifertest-pro`.

A. Romanelli, M. L. Lima, O. M. Quiroz Londoño, D. E. Martínez, and H. E. Massone. A gis-based assessment of groundwater suitability for irrigation purposes in flat areas of the wet pampa plain, Argentina. *Environmental management*, 50(3):490–503, 9 2012. ISSN 1432-1009. doi: 10.1007/s00267-012-9891-9. URL `http://www.ncbi.nlm.nih.gov/pubmed/22752689`.

R. Rossetto, I. Borsi, C. Schifani, E. Bonari, P. Mogorovich, and M. Primicerio. SID&GRID: integrating hydrological modeling in GIS environment. *Redondiconti della Società Geologica Italiana*, pages 282–283, 2013.

R. Rossetto, G. De Filippis, I. Borsi, L. Foglia, M. Cannata, R. Criollo, and E. Vázquez-Suñé. Integrating free and open source tools and distributed modelling codes in GIS environment for data-based groundwater management. *Environmental Modelling & Software*, 107:210–230, 9 2018. ISSN 13648152. doi: 10.1016/j.envsoft.2018.06.007. URL `https://linkinghub.elsevier.com/retrieve/pii/S1364815217309453https://www.sciencedirect.com/science/article/pii/S1364815217309453#!`

D. Rumbaugh and J. Rumbaugh. AquiferWin32 | pumping test analysis. *Environmental Software Online*, 1997. URL `http://www.groundwatersoftware.com/aquiferwin32.htm`.

R. R. Rumer. Resistance to flow through porous media. In *Flow through porous media.* Academic Press, 1969.

A. Rüping. Transform! Patterns for Data Migration. In J. Noble, R. Johnson, U. Zdun, and E. Wallingford, editors, *Transactions on Pattern Languages of Programming III. Lecture Notes in Computer Science*, chapter 7840, pages 1–23. Springer, Berlin, Heidelberg, 2013. ISBN 978-3-642-38675-6. doi: 10.1007/978-3-642-38676-3{_}1.

safe solutions. Safe Software | FME | Data Integration Platform, 2018. URL `https://www.safe.com/`.

X. Sánchez-Vila, J. Carrera, and J. P. Girardi. Scale effects in transmissivity. *Journal of Hydrology*, 183(1-2):1–22, 8 1996. ISSN 00221694. doi: 10.1016/S0022-1694(96)80031-X. URL `http://www.sciencedirect.com/science/article/pii/S002216949680031X`.

X. Sanchez-Vila, A. Guadagnini, and J. Carrera. Representative hydraulic conductivities in saturated groundwater flow. *Reviews of Geophysics*, 44(3):RG3002, 2006. ISSN 8755-1209. doi: 10.1029/2005RG000169. URL `http://www.agu.org/pubs/crossref/2006/2005RG000169.shtml`.

M. G. Schaap, F. J. Leij, and M. T. van Genuchten. Neural Network Analysis for hierarchical Prediction of Soil Hydraulic Properties. *Soil Science of America*, 62 (4):847–855, 1998.

M. G. Schaap, F. J. Leij, and M. T. van Genuchten. Rosetta: a Computer Program for Estimating Soil Hydraulic Parameters With Hierarchical Pedotransfer Functions. *Journal of Hydrology*, 251(3-4):163–176, 10 2001. ISSN 00221694. doi: 10.1016/S0022-1694(01)00466-8. URL `http://linkinghub.elsevier.com/retrieve/pii/S0022169401004668`.

R. Scheier. Data migration strategies and best practices, 2018. URL `https://searchstorage.techtarget.com/tip/Data-migration-strategies-and-best-practices`.

G. Schneebeli. *Hydraulique souterraine.* Eyrolles, 1966.

H. Schoeller. Les eaux souterraines. In Massoncie, editor, *Revue de géographie alpine*, chapter 4, pages 699–700. Massoncie, 1962.

B. Schönwalder. *Der Kutturtechniker.* 1928.

F. W. Schwartz and H. Zhang. *Fundamentals of ground water.* Wiley, 2003. ISBN 9780471137856.

M. Sen and T. Duffy. GeoSciML: Development of a generic GeoScience Markup Language. *Computers & Geosciences*, 31(9):1095–1103, 11 2005. ISSN 00983004. doi: 10.1016/j.cageo.2004.12.003. URL `http://dl.acm.org/citation.cfm?id=1650415.1650611`.

S. Serra and M. M. Marciszack. Methodology for construction of data migration processes in the context of developing systems. In *2018 13th Iberian Conference on Information Systems and Technologies (CISTI)*, pages 1–6. IEEE, 6 2018. ISBN 978-989-98434-8-6. doi: 10.23919/CISTI.2018.8399384. URL `https://ieeexplore.ieee.org/document/8399384/`.

A. Serrano Juan. *A hydrogeological approach in urban underground infrastructures.* PhD thesis, Universitat Politècnica de Catalunya (UPC), 7 2016. URL `http://www.tdx.cat/handle/10803/398402`.

A. Serrano-Juan and E. Vázquez-Suñé. Easy_QUIM. Chemical balances in water analysis and its graphic representation. Hydrogeology Group (CSIC-UPC). Spanish Council for Scientici Research (CSIC). Register number: CG3420743, 2015.

A. Serrano-Juan, R. Criollo, V. Velasco, C. Riera, and E. Vázquez-Suñé. FREE and Open Source Software Tools for WATer Resource Management FREEWAT User Manual-Volume 4 AkvaGIS (Hydrochemical Analysis Tools and Hydrogeological Analysis Tools). Technical report, FREEWAT, 2017.

R. G. Shepherd. Correlations Of Permeability And Grain-Size. *Ground Water*, 27(5): 633–638, 9 1989.

J.-W. Sjoukema, A. Bregt, and J. Crompvoets. Evolving Spatial Data Infrastructures and the Role of Adaptive Governance. *ISPRS International Journal of Geo-Information*, 6(8):254, 8 2017. ISSN 2220-9964. doi: 10.3390/ijgi6080254. URL `http://www.mdpi.com/2220-9964/6/8/254`.

C. S. Slichter. Theoretical investigations of the motion of ground waters. 19th Ann. Rept., pt.II-C. Technical report, U.S. Geological Survey, 1899.

Spatialite Development Team. Spatialite, 2011. URL `https://www.gaia-gis.it/fossil/libspatialite/home`.

J. M. Sperry and J. J. Peirce. A Model for Estimating the Hydraulic Conductivity of Granular Material Based on Grain Shape, Grain Size, and Porosity. *Ground water*, 33(6):892 – 898, 11 1995. doi: 10.1111/j.1745-6584.1995.tb00033.xAbout. URL `http://doi.wiley.com/10.1111/j.1745-6584.1995.tb00033.x`.

SPW-DGO3. Etat des nappes d'eau souterraine de Wallonie. Edition: Service public de Wallonie, DGO 3 (DGARNE), Belgique. Dépôt légal D/2017/11802/09. Technical report, Service public de Wallonie, 2016.

L. Steyaert and M. Goodchild. Integrating geographic information systems and environmental simulation models: a status review. In W. Michener and J. Brunt, editors, *Environmental information management and analysis: ecosystem to global scales.*, page 333–355. Taylor and Francis, London, UK, stafford s edition, 1994. URL `http://www.geog.ucsb.edu/~good/papers/211.pdf`.

G. Strassberg. *A geographic data model for groundwater systems.* PhD thesis, University of Texas. Austin, 2005. URL `http://hdl.handle.net/2152/2333`.

G. Strassberg, N. L. Jones, and D. R. Maidment. *ARC hydro groundwater: GIS for hydrogeology.* ESRI Press, 2011. ISBN 978-1589481985.

J. Sundell, L. Rosén, T. Norberg, and E. Haaf. A probabilistic approach to soil layer and bedrock-level modeling for risk assessment of groundwater drawdown induced land subsidence. *Engineering Geology*, 203:126–139, 3 2016. ISSN 0013-7952. doi: 10.1016/J.ENGGEO.2015.11.006. URL `https://www.sciencedirect.com/science/article/pii/S0013795215300910?via%3Dihub`.

Talend. Open Source ETL: Talend Open Studio for Data Integration, 2006. URL `https://www.talend.com/products/data-integration/data-integration-open-studio/`.

K. Tartar. Transportation Planning Division Spatial Data Management Workflow GIS-T 2008. In *INCOG*, Oklahoma, 2008. ISBN 918.584.7526. URL `http://www.gis-t.org/files/D86fn.pdf`.

A. C. Teodoro. *Frontiers in Information Systems. GIS : an overview of applications.* Bentham Science Publisher, 2018. ISBN 9781681086118. URL `https://ebooks.benthamscience.com/book/9781681086118/`.

K. Terzaghi. *Erdbaumechanik auf bodenphysikalischer grundlage.* F. Deuticke, Leipzig u. Wien, 1925.

K. Terzaghi and Wiley InterScience (Online service). *Theoretical soil mechanics.* J. Wiley and Sons, Inc., 1943. ISBN 9780470172766.

The MathWorks. MATLAB - The language of technical calculation - MATLAB & Simulink, 1984. URL `https://es.mathworks.com/products/matlab.html?s_tid=gn_loc_drop`.

A. K. Tiwari, A. K. Singh, A. K. Singh, and M. P. Singh. Hydrogeochemical analysis and evaluation of surface water quality of Pratapgarh district, Uttar Pradesh, India. *Applied Water Science*, 7(4):1609–1623, 7 2017. ISSN 2190-5487. doi: 10.1007/s13201-015-0313-z. URL `http://link.springer.com/10.1007/s13201-015-0313-z`.

K. Uma, B. Egboka, and K. Onuoha. New statistical grain-size method for evaluating the hydraulic conductivity of sandy aquifers. *Journal of Hydrology*, 108:343–366, 1 1989. doi: 10.1016/0022-1694(89)90293-X. URL `http://linkinghub.elsevier.com/retrieve/pii/002216948990293X`.

UPC. Ephebo, 2002.

E. Vázquez Suñé. *Urban Groundwater. Barcelona City Case Study.* PhD thesis, Universitat Politècnica de Catalunya (UPC), 10 2004. URL `http://www.tdx.cat/handle/10803/6232`.

E. Vázquez-Suñé, X. Sánchez-Vila, and J. Carrera. Introductory review of specific factors influencing urban groundwater, an emerging branch of hydrogeology, with reference to Barcelona, Spain. *Hydrogeology Journal*, 13(3):522–533, 8 2004. ISSN 1431-2174. doi: 10.1007/s10040-004-0360-2. URL `http://link.springer.com/10.1007/s10040-004-0360-2`.

E. Vázquez-Suñé, E. Abarca, J. Carrera, B. Capino, D. Gámez, M. Pool, T. Simó, F. Batlle, J. M. Niñerola, and X. Ibáñez. Groundwater modelling as a tool for the European Water Framework Directive (WFD) application: The Llobregat case. *Physics and Chemistry of the Earth*, 31(17):1015–1029, 2006. ISSN 14747065. doi: 10.1016/j.pce.2006.07.008.

E. Vázquez-Suñé, M. Ángel Marazuela, V. Velasco, M. Diviu, A. Pérez-Estaún, and J. Álvarez-Marrón. A geological model for the management of subsurface data in the urban environment of Barcelona and surrounding area. *Solid Earth*, 7(5):1317–1329, 9 2016a. ISSN 1869-9529. doi: 10.5194/se-7-1317-2016. URL `http://www.solid-earth.net/7/1317/2016/`.

E. Vázquez-Suñé, R. Criollo, I. Tubau, A. Jurado, I. Ansón, A. Soler, J. Carrera, M. Chesa, S. Burdons, R. Arandes, and M. Iglesias. Hidroquímica i contaminació dels aqüífers urbans de l'àrea de Barcelona. In *Geoquímica ambiental a Catalunya. Barcelona*, chapter 12. IGCC, 2016b. ISBN 9788439394716. URL `http://hdl.handle.net/10261/140967`.

V. Velasco. *GIS-based hydrogeologica platform for sedimentary media*. PhD thesis, Universitat Politècnica de Catalunya (UPC), 12 2013. URL `https://www.tdx.cat/handle/10803/135005`.

V. Velasco, P. Cabello, E. Vázquez-Suñé, M. López-Blanco, E. Ramos, and I. Tubau. A sequence stratigraphic based geological model for constraining hydrogeological modeling in the urbanized area of the Quaternary Besòs delta (NW Mediterranean coast, Spain). *Geologica Acta*, 10(4):373–393, 2012a. ISSN 16956133. doi: 10.1344/105.000001757.

V. Velasco, R. Gogu, E. Vázquez-Suñè, A. Garriga, E. Ramos, J. Riera, and M. Alcaraz. The use of GIS-based 3D geological tools to improve hydrogeological models of sedimentary media in an urban environment. *Environmental Earth Sciences*, 68 (8):2145–2162, 8 2012b. ISSN 1866-6280. doi: 10.1007/s12665-012-1898-2. URL `http://www.springerlink.com/index/10.1007/s12665-012-1898-2http://link.springer.com/10.1007/s12665-012-1898-2`.

V. Velasco, I. Tubau, E. Vázquez-Suñé, R. Gogu, D. Gaitanaru, M. Alcaraz, A. Serrano-Juan, D. Fernàndez-Garcia, T. Garrido, J. Fraile, and X. Sanchez-Vila. GIS - based hydrogeochemical analysis tools (QUIMET). *Computers & Geosciences*, 70:164–180, 5 2014. doi: 10.1016/j.cageo.2014.04.013. URL `http://dx.doi.org/10.1016/j.cageo.2014.04.013`.

V. V. Vesselinov, P. K. Mishra, and S. P. Neuman. Analytical Solution for Flow to a Partially Penetrating Well with Storage in a Confined Aquifer. In *American Geophysical Union, Fall Meeting 2009, abstract id. H43C-1047*, 2009. URL `http://adsabs.harvard.edu/abs/2009AGUFM.H43C1047V`.

T. Vienken and P. Dietrich. Field evaluation of methods for determining hydraulic conductivity from grain size data. *Journal of Hydrology*, 400(1-2):58–71, 3 2011. ISSN 00221694. doi: 10.1016/j.jhydrol.2011.01.022. URL `http://linkinghub.elsevier.com/retrieve/pii/S0022169411000473`.

M. Vukovic and A. Soro. *Determination of Hydraulic Conductivity of Porous Media from Grain-size*. Water Resources Publications, Littleton CO, 1992. ISBN 9780918334770.

B. Wagner, V. R. Tarnawski, V. Hennings, U. Muller, G. Wessolek, and R. Plagge. Evaluation of pedo-transfer functions for unsaturated soil hydraulic conductivity using an independent data set. *Geoderma*, 102(3-4):275–297, 8 2001.

M. Wagner and T. Wellhausen. Patterns for Data Migration Projects. Technical report, tngtech, 2011.

P. H. Walker, K. D. Woodyer, and J. Hutka. Particle-Size Measurements by Coulter Counter of Very Small Deposits and Low Suspended Sediment Concentrations in Streams. *Journal of Sedimentary Research*, 44(3):673–679, 9 1974. ISSN 1527-1404. doi: 10.1306/74D72AC7-2B21-11D7-8648000102C1865D. URL `http://jsedres.geoscienceworld.org/content/44/3/673.abstracthttp://archives.datapages.com/data/sepm/journals/v42-46/data/044/044003/0673.htm`.

L. Wang, C. R. Jackson, M. Pachocka, and A. Kingdon. A seamlessly coupled GIS and distributed groundwater flow model. *Environmental Modelling & Software*, 82:1–6, 2016. ISSN 13648152. doi: 10.1016/j.envsoft.2016.04.007.

K. Wende. A Model for Data Governance-Organising Accountabilities for Data Quality Management. In *ACIS*, 2007. URL `http://aisel.aisnet.org/acis2007/80`.

P. Wojda and S. Brouyère. An object-oriented hydrogeological data model for groundwater projects. *Environmental Modelling & Software*, 43:109–123, 5 2013. doi: 10.1016/j.envsoft.2013.01.015. URL `http://dx.doi.org/10.1016/j.envsoft.2013.01.015http://dx.doi.org/10.1016/j.envsoft.2013.01.015`.

R. B. S. Wojda, Piotr; Gogu. Conceptual model of hydrogeological information for a GIS-based Decision Support System in management of artificial recharge in semi-arid regions. In *XIth International Congress Université de Liège, Belgium*. International Association for Mathematical Geology, 2006.

K. Wortman. The Spatial Data Transfer Standard (FIPS 173): A Management Perspective. *Cartography and Geographic Information Systems*, 19(5):294–295, 1 1992. ISSN 1050-9844. doi: 10.1559/152304092783762227. URL `https://www.tandfonline.com/doi/full/10.1559/152304092783762227`.

J. Wösten. *Soil Quality for Crop Production and Ecosystem Health*, volume 25 of *Developments in Soil Science*. Elsevier, 1997. ISBN 9780444816610. doi: 10.1016/S0166-2481(97)80037-2. URL `http://dx.doi.org/10.1016/S0166-2481(97)80037-2`.

J. Wösten, Y. Pachepsky, and W. Rawls. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology*, 251(3-4):123–150, 10 2001. ISSN 00221694. doi: 10.1016/S0022-1694(01)00464-4. URL `http://linkinghub.elsevier.com/retrieve/pii/S0022169401004644`.

B. Wu, D. Lawless, J. Bisbal, J. Grimson, V. Wade, S. O' Donie, and R. Richardson. Legacy System Migration : A Legacy Data Migration Engine. In *Czechoslovak Computer Experts*, pages 129–138, 1997. URL `http://www.dtic.upf.edu/~jbisbal/publications/datasem97.pdf`.

Y. Yang and Y. Lin. Making the results of analysis accessible: ArcGIS Explorer plug-in aids natural resources management. ArcUser 14(1): 22–25., 2010. URL `http://www.esri.com/news/arcuser/0111/uwater.html`.

H. Ye, M. Brown, and J. Harding. GIS for All: Exploring the Barriers and Opportunities for Underexploited GIS Applications. *OSGeo Journal*, 13(1):19–28, 4 2013. ISSN 1994-1897. URL `https://journal.osgeo.org/index.php/journal/article/view/209`.

P. Yue and Z. Tan. 1.06 - GIS Databases and NoSQL Databases. In *Comprehensive Geographic Information Systems*, pages 50–79. Elsevier, 1 2018. ISBN 978-0-12-804793-4. doi: 10.1016/B978-0-12-409548-9.09596-8. URL `https://www.sciencedirect.com/science/article/pii/B9780124095489095968#s0165http://linkinghub.elsevier.com/retrieve/pii/B9780124095489095968`.

H. Zhan, L. V. Wang, and E. Park. On the horizontal-well pumping tests in anisotropic confined aquifers. *Journal of Hydrology*, 252(1-4):37–50, 10 2001. ISSN 00221694. doi: 10.1016/S0022-1694(01)00453-X. URL `http://www.sciencedirect.com/science/article/pii/S002216940100453X`.

Appendix A

Scientific production, projects and technical reports

Scientific manuscripts

Scientific publications directly related with this thesis are pointed below:

1. **Criollo, R.**; Velasco, V.; Vàzquez-Suñé, E.; Serrano-Juan, A.; Alcaraz, M.; García-Gil, A. (2015). An integrated GIS-based tool for Aquifer Test Analysis. *Environmental Earth Sciences.* doi: 10.1007/s12665-016-5292-3.
2. **Criollo, R.**; Vázquez-Suñé, E.; Velasco, V.; Nardi, A.; de Vries, L.M.; Riera, C.; Scheiber, L.; Jurado, A.; Brouyère, S.; Pujades, E.; Rossetto, R. (2018). AkvaGIS: An open source tool for water quality and quality management. *Computers & Geosciences.* doi: 10.1016/j.cageo.2018.10.012.
3. **Criollo, R.**; Vázquez-Suñé, E.; Cardona, F.; Burdons, S.; Enrich, M. (2019). On the hydrogeological data migration and integration. Submitted in *Environmental Modelling & Software.*
4. **Criollo, R.**; Vázquez-Suñé, E.; Sánchez-Vila, X.; Dentz, M.; Riera, J.; Velasco, V. (2019). On the reliability of hydraulic conductivity equations obtained by grain size data: new perspective for scalable applications. To be submitted in *Engineering Geology.*

Other scientific publications where the author had collaborated during this thesis:

1. Rossetto, R.; De Filippis, G.; Borsi, I.; Foglia, L.; Cannata, M.; **Criollo, R.**; Vázquez-Suñé, E.; Ghetta, M.; Rossetto, R. (2018). Integrating free and open source tools and distributed modelling codes in GIS environment for data-based groundwater management. *Environmental Modelling and Software.* doi: 10.1016/j.envsoft.2018.06.007.
2. Foglia, L.; Borsi, I.; Mehl, S.; De Filippis, G.; Cannata, M.; Vasquez-Sune, E.; **Criollo, R.**; Rossetto, R. (2018). FREEWAT, a Free and Open Source, GIS-Integrated, Hydrological Modeling Platform. *Groundwater.* doi: 10.1111/gwat.12654.
3. Serrano-Juan, A; Pujades, E.; Vázquez-Suñé, E.; Velasco, V.; **Criollo, R.**; Jurado, A. (2018). Integration of groundwater by-pass facilities in the bottom slab design for large underground structures. *Tunnelling and Underground Space Technology.* 71: 231-243. doi: 10.1016/j.tust.2017.07.020.
4. Rossetto, R.; De Filippis, G.; Borsi, I.; Foglia, L.; Cannata, M.; **Criollo, R.**; Vásquez-Suñé, E. (2017). Spatial analysis and simulation tools for groundwater management: the FREEWAT platform. *Acque Sotterranee – Italian Journal of Groundwater.* doi: 10.7343/as-2017-293.
5. Alcaraz, M.; Vázquez-Suñé, E.; Velasco, V.; **Criollo, R.** (2017). A loosely coupled GIS and hydrogeological modeling framework. *Environmental Earth Sciences.* doi: 10.1007/s12665-017-6709-3.
6. Tubau, I; Vázquez-Suñé, E.; Carrera, J.; Valhondo, C.; **Criollo, R.** (2017). Quantification of groundwater recharge in urban environments. *Science of the Total Environment.* doi: 10.1016/j.scitotenv.2017.03.118.
7. Serrano-Juan, A.; Vázquez-Suñè, E.; Monserrat, O.; Crosetto, M.; Hoffmann, C.; Ledesma, A.; **Criollo, R.**; Pujades, E.; Velasco, V.; Garcia-Gil, A.; Alcaraz, M. (2016). Gb-SAR interferometry displacement measurements during dewatering in construction works. Case of La Sagrera railway station in Barcelona, Spain. *Engineering Geology.* vol: 205. doi: 10.1016/j.enggeo.2016.02.014.

Chapters in books

1. Vázquez-Suñé, E.; **Criollo, R.**; Tubau, I.; Jurado, A.; Ansón, I.; Soler, A.; Carrera, J.; Chesa, M.J.; Burdons, S.; Arandes, R.; Iglesias, M. (2016).

Hidroquímica i contaminació dels aqüífers urbans de l'àrea de Barcelona. Geoquímica ambiental a Catalunya. Barcelona, 2016. Chapter 12. ISBN: 9788439394716.

Proceedings in congresses

1. De Filippis, G.; Borsi, I.; Foglia, L.; Cannata, M.; **Criollo, R.**; Vazquez-Suñe, E.; Rossetto, R. (2018). Simulating the hydrologic cycle in a GIS environment: present and future of the free and open source FREEWAT platform. *Session S03. Computational Methods in Water Resources (CMWR).* 3-7 June 2018. Saint-Malo, France. (Oral presentation).

2. **Criollo, R.**; Vázquez-Suné, E.; Riera, C.; Kukuric, N.; Borsi I.; Foglia L.; Cannata, M.; De Filippis, G.; Rossetto, R. (2018). Water management capacity building through the FREEWAT platform. *EGU 2018. Geophysical Research Abstracts.* Vol. 20, EGU2018-18956, 2018. EGU General Assembly, 8-13 April 2018. Wien, Austria. (Poster presentation).

3. **Criollo, R.**; Vázquez-Suné, E.; Burdons, S.; Enrich, M.; Varela, X. (2018). Urban groundwater quality. Update of the Barcelona city. *EGU 2018. Geophysical Research Abstracts.* Vol. 20, EGU2018-18937-1, 2018. EGU General Assembly, 8-13 April 2018. Wien, Austria. (**Highlighted** by session conveners as being of public interest:

   /EGU2018/orals/26645). (Oral presentation).

4. De Filippis, G.; Borsi I.; Foglia L.; Cannata, M.; **Criollo, R.**; Vázquez-Suné, E.; Kopač, I.; Panteleit, B.; Positano, P.; Nannucci, M.S.; Sapiano, M.; Svidzinska, D.; Grodzynskyi, M.; Rossetto, R. (2018). Joining participatory approach and spatially-based modelling tools for groundwater resource management. *EGU 2018. Geophysical Research Abstracts.* Vol. 20, EGU2018-8396. EGU General Assembly, 8-13 April 2018. Wien, Austria. (Oral presentation).

5. De Filippis, G.; Borsi I.; Foglia L.; Cannata, M.; **Criollo, R.**; Vázquez-Suné, E.; Kopač, I.; Panteleit, B.; Positano, P.; Nannucci, M.S.; Sapiano, M.; Svidzinska, D.; Grodzynskyi, M.; Rossetto, R. (2018). Promoting water resource management by means of spatially-based modelling tools and participatory approach. *XXVI Convegno Nazionale di Idraulica e Costruzioni Idrauliche.* 12-14 Settembre 2018. Ancona, Italy. (Poster presentation).

6. Foglia, L.; Mehl, S.; Rossetto, R.; Borsi, I.; Cannata, M.; De Filippis, G.; **Criollo, R.** (2017). The Horizon 2020 FREEWAT project: FREE and open source software tools for WATer resource management. *CWEMF 2017 Annual Meeting.* March 2017. Folsom, California, USA. (Oral presentation).
7. Foglia, L.; Borsi, I.; Cannata, M.; De Filippis, G.; **Criollo, R.**; Mehl, S.; Rossetto, R. (2017). How innovative ICT tools can enhance understanding of interactions between societal, hydrological and environmental changes (invited). *AGU127. Fall Meeting.* 11-15 December 2017. New Orleans, USA. (Oral presentation).
8. Foglia, L.; Rossetto, R.; Borsi, I.; Cannata, M.; **Criollo, R.** (2017). Using the new FREEWAT platform for water management: development and preliminary applications. *AHS2017-112. IAHS Scientific Assembly.* 10-14 July 2017. Port Elizabeth, South Africa. (Oral presentation).
9. Rossetto, R.; De Filippis, G.; Borsi, I.; Foglia, L.; Toegl, A.; Cannata, M.; Neumann, J.; Vázquez-Suñé, E.; **Criollo, R.** (2017). Achieving sustainable ground-water management by using GIS-integrated simulation tools: the EU H2020 FREEWAT platform. *EGU 2017. Geophysical Research Abstracts.* Vol. 19, EGU2017-18705. EGU General Assembly, 23-28 April 2017. Wien, Austria. (Poster presentation).
10. Marazuela, M.A.; Vázquez-Suñé, E.; **Criollo, R.**; García-Gil, A. (2017). The origin of the geothermal anomaly identified in the Barcelona underground (Spain): Future perspectives of this urban geothermal resource. *EGU 2017 Geophysical Research Abstracts.* Vol. 19, EGU2017-7059. EGU General Assembly, 23-28 April 2017. Wien, Austria. (Poster presentation).
11. Vázquez-Suñé, E.; **Criollo, R.**; Velasco, V.; Rosetto, R.; Borsi, I.; Foglia, L.; Canatta, M. (2017). FREE and open source software tools for WATer resource management. *WaterInnEU Final meeting.* 24-25 January 2017. Barcelona, Spain. (Oral presentation).
12. **Criollo, R.**; Velasco, V.; Vázquez-Suñé, E.; Nardi, A.; Marazuela, M.A.; Rosetto, R.; Borsi, I.; Foglia, L.; Canatta, M.; De Filippis, G. (2017). Open source GIS based tools to improve hydrochemical water resources management in EU H2020 FREEWAT platform. *EGU 2017 Geophysical Research Abstracts.* Vol. 19, EGU2017-19584. EGU General Assembly, 23-28 April 2017. Wien, Austria. (Poster presentation).

13. Rosetto, R.; **Criollo, R.**; Vázquez-Suñé, E.; Velasco, V. (2016). Open source and free software for water resource management: the H2020 FREEWAT platform (**invited**). *Water and Climate Day – COP22.* 9 November 2016. Marrakech, Morocco. (Oral presentation).

14. Velasco, V.; Vázquez-Suñé, E.; **Criollo, R.** (2016). FREE and open source software tools for WATer resource management. *Hidrogeología emergente, 50 Aniversario de CIHS 2016. Nuevos campos en la investigación y el desarrollo sostenible de las aguas subterráneas. FCIHS.* 13 May 2016. Barcelona, Spain. ISBN: 978-84-921469-3-2. (Poster presentation).

15. Serrano-Juan, A.; Vázquez-Suñé, E.; Pujades, E.; Velasco, V.; **Criollo, R.**; Jurado, A. (2016). Integración de medidas correctoras: efecto barrera y subpresión en un único diseño. Caso de La Sagrera (Barcelona, Spain). *Hidrogeología emergente, 50 Aniversario de CIHS 2016. Nuevos campos en la investigación y el desarrollo sostenible de las aguas subterráneas. FCIHS.* 13 May 2016. Barcelona, Spain. ISBN: 978-84-921469-3-2. (Poster presentation).

16. **Criollo, R.**; Velasco, V.; Vàzquez-Suñé, E.; Serrano-Juan, A.; Alcaraz, M.; García-Gil, A. (2016). Plataforma para la gestión de datos hidrogeológicos en entornos urbanos. *Hidrogeología emergente, 50 Aniversario de CIHS 2016. Nuevos campos en la investigación y el desarrollo sostenible de las aguas subterráneas. FCIHS.* 13 May 2016. Barcelona, Spain. ISBN: 978-84-921469-3-2. (Poster presentation).

17. **Criollo, R.**; Velasco, V.; Vázquez-Suñé, E.; Serrano, A.; Alcaraz, M.; Gacría-Gil, A. (2016). Aquifer characterization through an integrated GIS-based tool. *EGU 2016. Geophysical Research Abstracts.* Vol. 18, EGU2016-6445, 2016 EGU General Assembly, 17-22 April 2016. Wien, Austria. (Poster presentation).

18. Serrano, A.; Vázquez-Suñé, E.; Pujades, E.; Velasco, V.; **Criollo, R.;** Jurado, A. (2016). Infrastructure design integration to optimize structures and minimize groundwater impacts. Case of a bottom slab and groundwater by-pass integration in La Sagrera railway station, Spain. *EGU 2016. Geophysical Research Abstracts.* Vol. 18, EGU2016-964, 2016 EGU General Assembly, 17-22 April 2016. Wien, Austria. (Poster presentation).

19. Serrano, A.; Vázquez-Suñé, E.; Monserrat, O.; Crosetto, M.; Hoffmann, C.; Ledesma, A.; **Criollo, R.**; Pujades, E.; Velasco, V.; Garcia-Gil, A.; Garcia-Alcaraz, M. (2015). GB-SAR Interferometry for Structure Monitoring Con-

struction Works. *AGU Fall Meeting.* 14-18 December 2015. San Francisco, USA. (Poster presentation).

20. Velasco, V.; Vázquez-Suñé, E.; **Criollo, R.**; Alcaraz, M.; Serrano, A.; García, A. (2015). GIS-based analysis tools for facilitating the development of hydrogeological models. *42nd IAH International Congress, AQUA2015.* 13-18 September 2015. Rome, Italy.
21. Velasco, V.; Alcaraz, M.; Vázquez-Suñé, E.; **Criollo, R.**; Serrano, A.; García, A. (2015). GIS-based tools for facilitating the application of the groundwater related directives. *The Seventh International Conference on Advanced Geographic Information Systems, Applications, and Services. GEOProcessing 2015.* 22-27 February 2015. Lisbon, Portugal. (Poster presentation).

Software Registration / Patents

The strategic relevance of hydrological and hydrogeochemical of our research is in many cases highly applied and focused into key areas in environment and society such as groundwater resources and the management of aquifers. This relevance is reflected in the recent patent registrations.

1. MJ-Pumpit. **Criollo, R.**; Serrano-Juan, A.; Vázquez-Suñé, E. (2015). Registered by the Spanish Council for Scientific Research (CSIC). Register number: CG3420754.

   MJ-PUMPIT is a GUI for the MariaJ code (Carbonell et al., 1997) for pumping test interpretation. MJ-Pumpit simplifies data handle and the model selection required to execute the code and makes it possible to apply diagnostic plots and create other pumping tests interpretation methods through the use of the tools offered by MS Excel.

2. HYYH. **Criollo, R.**; Serrano-Juan, A.; Vázquez-Suñé, E.; Velasco, V.; Alcaraz, M. (2015). Registered by the Spanish Council for Scientific Research (CSIC). Register number: DW7637576.

   HYYH is a GIS-based toolbar developed in ArcGIS (ESRI). Their tools help users to manage, visualise and analyse hydrogeological data. Additionally, it can automatically export pumping test data to MJ-PUMPIT for further analyses.

HYYH and MJ-Pumpit have been implemented as teaching subject of the professional master's degree of the "Fundación Internacional Centro de Hidrologia Subterránea" (FCIHS), recognised by UNESCO. Within the framework of this teaching subject, these tools have been applied in several master theses and more than 50 students have attended the subject where they have learned the use of GIS tools to facilitate the environmental assessment of the subsoil. Recently, MJ-Pumpit and HYYH have been included in the curricula of courses entitled "hydrogeologie quantitative" and "captage des eaux souterraines" at the University of University Peleforo Gon Coulibaly of Korhogo (Côte d'Ivoire).

Finally, these tools are being applied in widely environmental projects such as groundwater management, civil works, mining, among others. These applications have been useful to fix some bugs, to improve and to maintain them in latest versions of ArcGIS.

Projects involved and technical reports

1. FREE and open source software tools for WATer resource management. FREEWAT project. (Horizon 2020 agreement number: 642224). (IDAEA-CSIC, 2015 - 2017). Member of the steering group and WP3 leader from July 2016 until end of the project.
2. NAVIA. Plataforma para la gestión de datos hidrogeológicos a través de aplicación móvil (design and preparation of the proposal of the project). (IDAEA-CSIC and BCASA, June 2018).
3. Secció d'aigües subterrànies del Pla tècnic per a l'aprofitament dels recursos hídrics alternatius a Barcelona (PLARHAB). Edició 2017. (BCASA, February-June 2018).
4. Requeriments de la implementació de la plataforma HYDORGIS als Sistemes de BCASA (BCASA, March 2015).
5. Correspondència de camps entre bases de dades HYDOR i SITE. (BCASA, 2016 - 2018).
6. Manual de migració i integració de dades dels Sistemes de BCASA a HYDOR (BCASA, 2018).
7. Modelización numérica hidrogeológica del proyecto minero Volcan. Ticlio. Perú (IDAEA-CSIC and WES, September 2016).

8. Elaboración de la fase I del diseño y formulación de acciones y estrategias para la evaluación y gestión integrada del agua subterránea en dos áreas priorizadas de la Orinoquia Colombiana. Ministerio de ambiente y desarrollo sostenible. Colombia (2015).
9. Avaluació de les possibilitats d'esgotament del freàtic durant la construcció de la nova plaça de les Glòries (IDAEA-CSIC and BCASA, February - April 2015).
10. Seguimiento hidrogeológico de los procesos de agotamiento del freático en el ámbito de la ejecución del túnel de la plaza de las Glorias (IDAEA-CSIC and BCASA, September 2015 - September 2016).
11. Estudio y seguimiento hidrogeológico de las obras de la línea de alta velocidad Madrid-Barcelona-Figueres, correspondientes a la construcción de la estación de la Sagrera y sus accesos (IDAEA-CSIC and UTE Sagrera, March 2014 - December 2014).
12. Estudi i seguiment de l'evolució del drenatge i dels nivells freàtics a l'entorn de la plaça de la Vila de Sant Adrià del Besòs. Període 2013-2016 (2016).
13. Mezcla y dispersión en el transporte de energía y solutos. MEDISTRAES project. (CICYT agreement number: CGL2013-48869-C2-1-R) (IDAEA-CSIC, 2014 - 2016).
14. Evaluación de las posibilidades de aprovechamiento de la anomalía geotérmica detectada en el sector de Fondo, Santa Coloma de Gramenet (IDAEA-CSIC and Consorci Besòs - Santa Coloma de Gramenet Council, October 2014 - March 2015).
15. Modelación del acuífero de Calama, sector medio de la Cuenca del río Loa, Antofagasta. (IDAEA-CSIC and Matraz Consultores, 2012 - 2013).
16. Predicting and monitoring the long term behaviour of CO2 injected in deep geological formations. PANACEA project. (FP7 agreement number: 282900). (2012 - 2014).
17. Estudio del posible efecto de los lodos bentoníticos previstos para el pilotaje del viaducto del Vial Puerto-Aeropuerto, Barcelona (IDAEA-CSIC, November 2012).

Appendix B

Batch file to execute FME frameworks with time flags

```
:: ARXIU PER A EXECUTAR TOTS ELS .FMW DEL TRASPAS
:: A HYDORGIS
:: OJO: ‘PRÈVIAMENT CAL TENIR EL SW ENGEGAT EN
:: ’fme importador & exportador’
:: I CAL ACTUALITZAR LES VARIABLES DE
:: DATA_INI, DATA_FI I DATA_INI_MDB
:: PER DEIXAR EL RANG DE DATES QUE ENS INTERESSA
:: DATA_INI1 I DATA_INI_MDB HAURÀN DE SER IGUALS
:: PER FER LA COMPARACIÓ
:: DES DEL TEMPS INICIAL DE CÀRREGA DE DADES

ECHO OFF

ECHO -------------------------------------------
ECHO - DATES INCREMENTALS NIVELLS PIEZOMETRICS (HG) -

SET HG_DATA_INI="^’01/05/2018^’"
SET HG_DATA_FI="^’03/07/2018^’"
SET HG_DATA_INI_MDB="^#01-05-2018#^"
```

Batch file to execute FME frameworks with time flags

```
ECHO -------------------------------------------
ECHO - DATES INCREMENTALS QUALITAT AIGUA (HQ) -

SET HQ_DATA_INI="^’1/04/2018^’"
SET HQ_DATA_FI="^’3/07/2018^’"
SET HQ_DATA_INI_MDB="^#01-04-2018#^"

ECHO -------------------------------------------
ECHO - Obrir relacions a la base se dades HYDOR -
fme 00_HYDOR-Open_Relations.fmw

ECHO -------------------------------------------
ECHO - Inici de posicions (SITE) -
fme 01_SITE-HYDOR.fmw

ECHO -------------------------------------------
ECHO - Inici de Mines i Rieres -
fme 04_SITE-SHPMines-_Rieres.fmw

ECHO -------------------------------------------
ECHO - Inici de Hidroquimica -
ECHO Data Inicial: "%HQ_DATA_INI%"
ECHO Data Final: "%HQ_DATA_FI%"
ECHO Data Inicial Comparacio: "%HQ_DATA_INI_MDB%"
fme 03_BDECA-HidroquimicaHYDOR.fmw
    --Data_Inici "%HQ_DATA_INI%" ^
    --Data_Final "%HQ_DATA_FI%" ^
    --Data_Ini_MDB "%HQ_DATA_INI_MDB%"

ECHO -------------------------------------------
ECHO - Inici de Hidrogeologia Carregar Nivells -
ECHO Data Inicial: "%HG_DATA_INI%"
ECHO Data Final: "%HG_DATA_FI%"
ECHO Data Inicial Comparacio: "%HG_DATA_INI_MDB%"
fme 02_BDE-NivellsHYDOR_1LoadNivells.fmw
    --Data_Inici "%HG_DATA_INI%" ^
```

```
    --Data_Final "%HG_DATA_FI%" ^
    --Data_Ini_MDB "%HG_DATA_INI_MDB%"

ECHO ------------------------------------------
ECHO - Inici de Hidrogeologia Esborrar Anomalies -
fme 02_BDE-NivellsHYDOR_2DeleteAnom.fmw

ECHO ------------------------------------------
ECHO - crear relacions a la base se dades HYDOR -
fme 05_HYDOR-Close_Relations.fmw

ECHO ------------------------------------------
ECHO ------------------------------------------
ECHO Finalitzat el traspas de dades cap a HYDOR.
ECHO Si el traspas ha donat errors, si us plau,
ECHO reviseu els arxius *.log

pause
```

Appendix C

Detailed information of the major methods based on grain size data

Specific nomenclature used in the evaluation of *K* by grain size methods compiled

K: hydraulic conductivity

K_0: intrinsic permeability

d_e: equivalent diameter

m: porosity

m(%): porosity in percentage

θ: fluid temperature [ºC]

g: gravity acceleration

υ: fluid dynamic viscosity

μ: fluid kinematic viscosity [Pa · s]

γ: fluid specific weight [N/m^3]

Detailed information of the major methods based on grain size data

P_n: percentages keep between two consecutive sieves (Fair and Hatch method)

N: number of P_n (Fair and Hatch method)

σ: grain size log-normal deviation

s: sand percentage contained

c: clay percentage contained

u: silt percentage contained

ρ_b: bulk density [g/cm^3]

ρ_w: fluid density [g/cm^3]

om: organic matter percentage contained

CP: capillary porosity (Li et al., 2008)

C_u: coefficient of uniformity, $C_u = d_{60}/d_{10}$

Φ: diameter of particles

n: equal fractions of the total pore space

r_i: (in cm) is the mean radius of the r^{th} fraction. r decrease for r_1 to r_n

GMPS: geometric mean of the particle size (in mm)

GSD: geometric standard deviation of the particle sizes

f_i: fraction of particles between two consecutive sieves (Choo et al., 2016)

d_{Li}: larger sieve size between two consecutive sieves (Choo et al., 2016)

d_{Si}: smaller sieve size between two consecutive sieves (Choo et al., 2016)

a: cementation constant (other equations use a for other parameters but each of them are specified in each method).

t: fitting parameter for saturation (t $\approx$ 2, typically)

v: Archie's m exponent, which is also called as shape factor or cementation factor (called m in the original source) (Archie, 1942)

ρ: electrical resistivity. It is equal to $1/\sigma$ (where σ_{mix}, σ_p,and σ_w are the electrical conductivities of the media (or mixture), particles, and pore water, respectively (S/m)

g_n: weight of each specific diameter of particles [g] (Schönwalder, 1928)

d_n: diameter of each weight of each specific diameter [mm] (Schönwalder, 1928)

Specific parameters found in the literature

Hazen

C use to be C = 100 (Custodio and Llamas, 1984), but it could be also:

41 - 146 (Taylor, 1948, in Carrier, 2003)

100 - 150 (Leonards, 1962, in Carrier, 2003)

100 - 1000 (Mansur and Kaufman, 1962, in Carrier, 2003)

100 - 150 (Terzaghi and Peck, 1964)

90 - 120 (Cedergren, 1967, in Carrier, 2003)

1 - 42 (Lambe and Whitman, 1969)

45 - 140 (Bear, 1979, in Carrier, 2003)

40 - 120 (Holtz and Kovacs, 1981, in Carrier, 2003)

100 - 150 (Das, 1997, in Carrier, 2003)

100 - 150 : (Domenico and Schwartz, 1998)

80 - 120 (Coduto, 1999, in Carrier, 2003)

Tison (in Schoeller, 1962) proposed:

$$C = 150 \cdot \left(\frac{m}{0.45}\right)^6$$

Slichter

n≈3,3 (Custodio and Llamas, 1984)

Castany (1963), in Custodio and Llamas, 1984 described Slichter equation as:

$$C_{slichter} = \frac{100}{58053 \cdot m(\%)^{-3.415} \cdot m^{3.3}}$$

Terzaghi

α = 800 to rounded grains and α= 460 to irregular shape (Custodio and Llamas, 1984)

Kozeny

Detailed information of the major methods based on grain size data

Schoeller, (1962) describes this equation as follow:

$$S_* = \frac{6 \cdot (1-m)}{d_e}$$

Schönwalder and Zunker (for both methods)

$$S_s = \frac{1}{dw}$$

Where dw is the specific diameter [mm]:

$$\frac{1}{dw} = \frac{g_1}{d_1} + \frac{g_2}{d_2} + \frac{g_3}{d_3} + \cdots = \sum \frac{g_n}{d_n}$$

Where g_i is the specific weight of the g_n sample and d_i are their the specific diameter.

But d_w used to be $d_w \approx$ d_e (Schoeller, 1962).

Fair-Hatch

P_n is the percentages keep between two consecutive sieves; Dn is the mean diameter between two sieves [m]; N is the number of P_n

$B \approx$ 5 and $\alpha = 6$ for spherical grains and $\alpha = 7,7$ for angular grains, respectively (Batu, 1998).

Bakhmeteff-Feodoroff

c= 640 (Schoeller, 1962).

Krumbein-Monk

b= 710 (Schwarz and Zhang, 2003).

Irmay

m_0= [0.2, 0.5] (Schoeller, 1962).

Kozeny-Carman

d_m is the representative diameter; $d_m \approx d_e$ (Batu, 1998).

Marshall

r_i is the main radius for each fraction [cm] (Cronican and Gribb, 2004; Marshall, 1958).

Kozeny, Fair and Hatch

S/V represents the specific surface

$$\left(\frac{S}{V}\right)^2 = \sum \frac{\eta}{d^2}$$

Where η is the percent keep between two consecutive sieves and d is the geometric median of the sieve size (Fair and Hatch, 1933, in Custodio and Llamas, 1984).

Nevertheless, there are other option proposed proposed by Bear et al. (1933, in Custodio and Llamas, 1984):

$$c \cdot \left(\frac{V}{S}\right)^2 = A \cdot {d_e}^2$$

Where: A= [1/150 - 1/175].

Rumer

β is the shape parameter. Its value is $\Pi/6$ in case of spherical grains.

α = [152- 207] is for non-spherical to spherical grains (in Custodio and Llamas, 1984).

Bloemen

GSI is the grain size distribution index (Bloemen, 1980, in Li et al., 2008).

Shepherd

Developed for channel sediments. Parameters a and b are empiric parameters within the next intervals: a = [1014 - 208808] gpd/ft2 and b = [1.11 - 2.05] (Shepherd, 1989).

Ahuja

Φe: total porosity minus the volumetric water content at 33kPa of suction. (Cronican and Gribb, 2004).

MacDonald

M1 and M2 are the first and second moment from the size grain distribution [mm]. (Li et al., 2008; MacDonald et al., 1991).

Alyamani and Sen

I_0: is the value corresponding with the intersection between the line that join d_{10} and d_{50} from the granulometric size distribution and the abscise axis (Alyamani and Sen, 1993; Sezer et al., 2009).

Rawls and Brakensiek

Application domain is advice assuming 1.5% of organic matter (Cronican and Gribb, 2004).

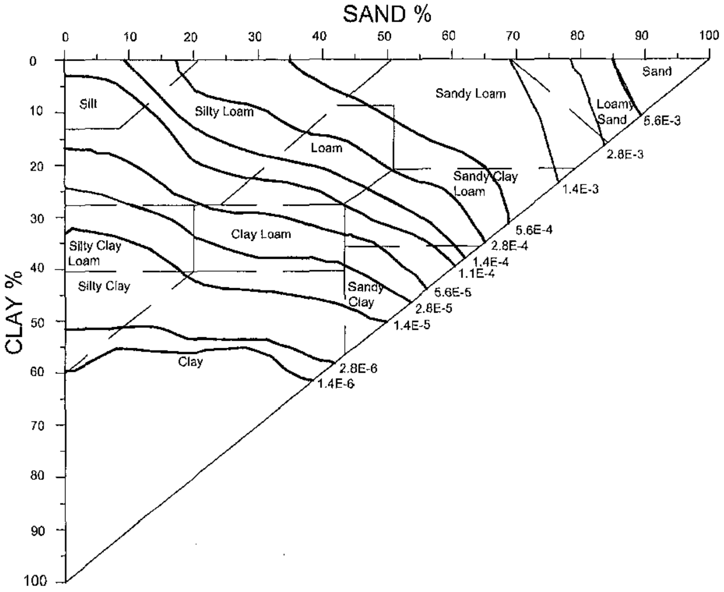


Fig. C.1 Percentage of sand and clay (from Cronican and Gribb, 2004).

Sperry and Pierce

γ: angle of repose in degrees (Sperry and Peirce, 1995).

Barr

C_c is the shape parameter from [5 ± 0.25]; C_s is the area surface parameter ([1.1 ± 0.1]) and S_0 is the area surface per solid mass unit (Barr, 2001).

Appendix D

Equations for permeability estimation by grain size and other material properties

Equations for permeability estimation by grain size and other material properties

Methods	Equation	Comments	Original source	Other sources
Hazen	$K = C \cdot d_e^2$ $K = \left(\frac{g}{\nu}\right) \cdot 6 \cdot 10^{-4} \cdot [1 + 10 \cdot (m - 0.26)] \cdot d_e^2$	K $[cm/s]$ d_e $[mm]$ Based in poorly grade sands, packed at medium density $0.1mm < d_e < 3mm$ Uniform coef. (Cu) $Cu < 5$ C = 100	Hazen, 1892	Terzaghi & Peck, 1964 Uma et al., 19889- Kasenow, 2002 Carrier, 2003 Odong, 2007 Sezer, 2009 Chapuis, 2012
Hazen with Taylor	$K = 1.157 \cdot (d_e)^2 \cdot [0.70 + 0.03 \cdot (T)]$	K $[cm/s]$ d_e $[mm]$ T [$\circ$C] For loose uniform sand $e \sim e_{max}$	Taylor, 1948	Taylor, 1948 Chapuis, 2012
Hazen with Tison	$K = 150 \cdot \left(\frac{m}{0.45}\right)^6 \cdot d_e^2$	K $[cm/s]$ d_e $[mm]$ Not recommended for $m < 0.4$	Tison, 1955	Schoeller, 1962

Methods	Equation	Comments	Original source	Other sources
Hazen with Schoeller	$K = 150 \cdot \left(\frac{m}{0.45}\right)^6 \cdot (d_e)^2 \cdot \left[\frac{0.70 + 0.03 \cdot T}{86400}\right]$	K $[cm/s]$ d_e $[mm]$ T $[\circ\mathrm{C}]$	Schoeller, 1962	
Sullivan & Hertel	$K = \left(\frac{\gamma}{\mu}\right) \cdot C \cdot d_{10}^2 \cdot \left(\frac{m^3}{(1-m)^2}\right)$	C (0.01) depends on the shape of the grains and the structure of the soil * Not applied in this study	Sullivan & Hertel, 1942	Schoeller, 1962
Schönwalder	$K = \left(\frac{\gamma}{\mu}\right) \cdot 0.0137 \cdot \left(\frac{m}{1-m}\right) \cdot \left(\frac{1}{S_s^2}\right)$ $where:$ $S_s = \frac{1}{d_w}$ $\frac{1}{d_w} = \sum \frac{g_n}{d_n}$	K $[cm/s]$ d_{10} $[cm]$ S_s specific surface of the sand] d_{10} can be applied by d_w	Schnwalder, 1928	Schoeller, 1962

Methods	Equation	Comments	Original source	Other sources
Terzaghi	$K = \alpha \cdot \left(m - 0.13 \cdot / \sqrt[3]{1-m}\right)^2 \cdot d_e^2$ $K = \left(\frac{g}{\nu}\right) \cdot C_t \cdot \left(\frac{m - 0.13}{\sqrt[3]{1-m}}\right)^2 \cdot d_e^2$	Based in large grain size of sand $d_e = d_{10}$ (in mm) $0.0061 < C_t < 0.00107$(typically) In this study $C_t = 0.0084$ (from Sezer, 2009)	Terzaghi, 1925	Custodio & Llamas, 1984 Schoeller, 1962 Cheng & Chen, 2007 Odong, 2007 Sezer, 2009 Chapuis, 2012
Slichter	$K = c \cdot d_e^2 \cdot m^n$ $k = \left(\frac{g}{\nu}\right) \cdot n^{3.287} \cdot d_e^2$ $K = \left(\frac{g}{\nu}\right) \cdot 1 \cdot 10^{-2} \cdot n^{3.287} \cdot d_e^2$	K $[cm/s]$ k intrinsic permeability $[cm^2]$ $0.01mm < d_e < 5mm$ (From Vuković & Soro, 1992)	Slichter, 1899	Custodio & Llamas, 1984 Vuković & Soro, 1992 Kasenow, 2002 Odong, 2007 Sezer, 2009
Zunker	$K = \left(\frac{\gamma}{\mu}\right) \cdot a \cdot \left(\frac{m_d}{1-m}\right) \cdot \left(\frac{1}{d_{10}^2}\right)$	Based in cases of fine and medium grain size sand m_d is the dynamic effective porosity * Not applied in this study	Zunker, 1930	Schoeller, 1962 Kasenow, 2002

Methods	Equation	Comments	Original source	Other sources
Hullvert & Feben	$K = (29.2 + T) \cdot \left(\frac{0.49554 \cdot d_s^{1.85}}{0.6943 - m} \right)$	* Not applied in this study	Hullvert & Feben, 1933	Schoeller, 1962
Fair & Hatch	$K = \left(\frac{\gamma}{\mu} \right) \cdot \left[\frac{m^3}{(1-m)^2} \right] \cdot \left[\frac{1}{\beta \cdot \left(\frac{\alpha}{100} \cdot \sum_{n=0}^{N} \frac{p_n}{d_n} \right)^2} \right]$	Based in cases of fine and medium grain size sand β is a packing factor (approx. 5) a is a sand shape factor (from 6 to 7.7 for spherical grains to angular grains respectively)	Fair & Hatch, 1933	Batu, 1998
Marshall	$k = \frac{1}{8} \cdot [m^2 \cdot n^2 \cdot (r_1^2 + 3r_2^2 + 5r_3^2 + ... + (2m-1) \cdot r_n^2)]$	For isotropic materials which their mean radius of pores for each of n equal fractions of the total pore space are represented by the corresponding mean radii $(r1, r2, ..., r_n)$ * Not applied in this study	Marshall, 1958	Cronican & Gribb, 2004; 2007

Methods	Equation	Comments	Original source	Other sources
Harleman	$k = 6.54 \cdot 10^{-4} \cdot d_{10}^2$ $K = \left(\frac{\gamma}{\mu}\right) \cdot 6.54 \cdot 10^{-4} \cdot d_{10}^2$	k is the intrinsic permeability For uniformly grade sand*	Harleman, 1963	Schwarz & Zhang, 2003 Uma et al., 19889*
Kozeny, Fair & Hatch	$k = c \cdot \left(\frac{V}{S}\right)^2 \cdot \left(\frac{m^3}{(1-m)^2}\right)$	k is the intrinsic permeability where $\left(\frac{V}{S}\right)^2 = \sum \frac{\eta}{d^2}$ η percentage retained between two consecutive sieves d is the geometric mean of the sieve sizes used Noticed that $c\left(\frac{V}{S})^2\right) = Ad_e^2$ where A ranges between 1/150 and 1/175 * Not applied in this study	Fair & Hatch, 1933	Custodio & Llamas, 1984 Schneebeli, 1966
Rumer	$k = \left(\frac{\beta \cdot m^2}{\alpha \cdot (1-m)}\right) \cdot d^2$	k is the intrinsic permeability $d \simeq d_{10}$ β is the shape factor ($\beta = \pi/6$ for spheres) α is an experimental factor (ranges between 152-207 for spheres) If the granulometry is not homogeneous, $\alpha = d_{60}/d_{10}$	Rumer, 1969	Custodio & Llamas, 1984

Methods	Equation	Comments	Original source	Other sources
Krumbein & Monk	$k = b \cdot d^2 \cdot e^{(-1.31 \cdot \sigma)}$	Based in unconsolidated sands k intrinsic permeability [darcys] $d \simeq d_{10}$ $m(\%) = 40 \pm 5$ $0 < \sigma < 0.5$	Sullivan & Hertel, 1942	Schwarz & Zhang, 2003 Cronican & Gribb, 2004; 2007
Irmay	$K = \left(\frac{\gamma}{\mu}\right) \cdot C \cdot d_{10}^2 \cdot \left(\frac{(m - m_0)^3}{(1-m)^2}\right)$	m_0 Irmay porosity factor (0.2 - 0.5 m.) $C = X0.01$, where X is function of shape and package * Not applied in this study	Irmay, 1954; 1956	Schoeller, 1962
Kozeny-Carman	$K = \left(\frac{\gamma}{\mu}\right) \cdot \left(\frac{m^3}{(1-m)^2}\right) \cdot \left(\frac{d_m^2}{180}\right)$ $K = \left(\frac{g}{\nu}\right) \cdot 8.3e^{-3} \cdot \left(\frac{m^3}{(1-m)^2}\right) \cdot d_m^2$ $K = C \cdot \frac{g}{\mu_w \cdot \sigma_w} \cdot \frac{e^3}{S_s^2 \cdot G_s^2 \cdot (1+e)}$	It is not appropriate for either soil with effective size above 3mm of for clavey soils $d_m \simeq d_{10}$ C depends on the porous space geometry	Carman, 1938 Kozeny & Carman, 1956	Kozeny, 1953 Köhler, 1965 Bear, 1972 Vuković & Soro, 1992 Batu, 1998 Carrier, 2003 Odong, 2007 Sezer, 2009 Chapuis, 2012 Choo et al., 2016

Methods	Equation	Comments	Original source	Other sources
Brakensiek	$K = 24 \cdot e^x$ $where:$ $x = 19.52348 \cdot m - 8.96847 - 0.028212 \cdot c$ $+0.00018107 \cdot s^2 - 0.0094125 \cdot c^2 - 8.395215 \cdot m^2$ $+0.077718 \cdot s \cdot m - 0.00298 \cdot s^2 - 0.019492 \cdot c^2 \cdot m^2$ $+0.0000173 \cdot s^2 \cdot c + 0.02733 \cdot c^2 \cdot m + 0.001434 \cdot s^2$ $\cdot m - 0.0000035 \cdot c^2 \cdot s$	K $[cm/day]$ s content of sand particles (in %) where $\phi \geq 50\mu$m c content of clay particles (in %) where $\phi \leq 2\mu$m ϕ represents the diameter	Brakensiek et al., 1984	Li et al., 2008
Cosby	$K = 60.96 \cdot 10^{-0.6+0.0126 \cdot s - 0.0064 \cdot c}$	K [cm/day] s content of sand particles (in %) where $\phi \geq 50\mu$m c content of clay particles (in %) where $\phi \leq 2\mu$m ϕ represents the diameter	Cosby et al., 1984	Li et al., 2008
Campbell	$K = 339 \cdot \left(\frac{1.3}{\rho_b}\right)^2 \cdot e^{(-6.9 \cdot c - 3.7 \cdot u)}$ $where:$ $b = GMPS^{-0.5} + 0.2 \cdot GSD$	K [cm/day] where $\phi \geq 50\mu$m c content of clay particles (in %) where $\phi \leq 2\mu$m u content of silt particles (in %) where $50 > \phi > 2\mu$m ϕ represents the diameter	Campbell, 1985	Li et al., 2008 (1)

Methods	Equation	Comments	Original source	Other sources
Krüger	$K = \left(\frac{g}{\nu}\right) \cdot C_K \cdot \frac{m}{(1-m)^2} \cdot d_e^2$ $where:$ $d_e = \sum_{i=1}^{i=n} \triangle g_i \cdot \frac{2}{(d_i^g + d_i^d)}$	K $[m/day]$ d_e $[mm]$ Most suitable for medium grain size sands $C_K = 4.35 \cdot 10^{-5}$		Vuković & Soro, 1992 (10)
Puckett	$K = 4.36 \cdot 10^{-3} \cdot e^{(-0.1975 \cdot c)}$	K [cm/s] c content of clay particles (in %) where $\phi \leq 2\mu$m Can be applied when: $34.6\% < s < 88.5\%$ $1.4\% < c < 42.1\%$	Puckett et al., 1985	Cronican & Gribb, 2004; 2007
Saxton	$K = 24 \cdot e^x$ $where:$ $x = 12.012 - 7.55 \cdot 10^{-2} \cdot s +$ $\frac{-3.895 + 3.671 \cdot 10^{-2} \cdot s - 0.1103 \cdot c + 8.7546 \cdot 10^{-4} \cdot c^2}{0.332 - 7.251 \cdot 10^{-4} \cdot s + 0.1276 \cdot log_{10}(c)}$	K [cm/day] c content of clay particles (in %) where $\phi \leq 2\mu$m s content of sand particles (in %) where $\phi \geq 50\mu$m Can be applied when: $1 < \rho_b < 1.8[g/cm^3]$ $om > 8\%$ $c < 6\%$	Saxton et al., 1986	Li et al., 2008 (2)

Methods	Equation	Comments	Original source	Other sources
Shepherd	$K = a \cdot d^b$	K $[gpd/ft^2]$ a between 1014 - 208808 $[gpd/ft^2]$ b between 1.11 - 2.05	Shepherd, 1989	Batu, 1998
Sauerbrei	$K = 3.49 \cdot \frac{m^3}{(1-m)^2} \cdot \tau \cdot d_e^2$	K $[m/s]$ d_e $[mm]$ Most suitable for sand and sandy clay $d_e < 0.5mm$ $\tau \sim 1.05 \cdot 10^{-6} \cdot \frac{1}{\nu}$ $\tau = 1.052$ for temp. $= 20°$ C		Vuković & Soro, 1992 (10) Kasenow, 2002
Rawls & Brakensiek	$K = e^x$ $where:$ $x = 19.52348 \cdot m - 8.96847 - 0.028212 \cdot c$ $+0.00018107 \cdot s^2 - 0.0094125 \cdot c^2$ $-8.395215 \cdot m^2 + 0.077718 \cdot s \cdot m$ $-0.00298 \cdot s^2 \cdot m^2 - 0.019492 \cdot c^2 \cdot m^2$ $+0.0000173 \cdot s^2 \cdot c + 0.02733 \cdot c^2 \cdot m$ $+0.001434 \cdot s^2 \cdot m - 0.0000035 \cdot c^2 \cdot s$	K $[cm/hours]$ Can be applied when: $5\% < s < 70\%$ $5\% < c < 60\%$	Rawls & Brakensiek, 1989	Cronican & Gribb, 2004; 2007

Methods	Equation	Comments	Original source	Other sources
Ahuja	$K = 1058.4 \cdot \phi_e^{3.3545}$	K $[cm/hours]$ ϕ_e is the total volumetric porosity without water content for a 33 kPa of suction * Not applied in this study	Ahuja et al., 1989	Cronican & Gribb, 2004; 2007
McDonald	$K = \frac{1920}{\pi^2 \cdot \mu} \cdot \frac{m^3}{(1-m)^2} \cdot \left(\frac{M_2}{M_1}\right)^2$	K $[cm/day]$ M_1 and M_2 are the first and the second momentum of the grain size distribution (in mm) * Not applied in this study	McDonald et al., 1991	Li et al., 2008
Dane & Puckett	$K = 8.44 \cdot 10^{-5} \cdot e^{(-0.144c)}$	K $[cm/s]$ c content of clay particles (in %) where $\phi \leq 2\mu$m	Dane & Puckett, 1992	Cronican & Gribb, 2004; 2007
Jabro	$log(K) = 9.56 - 0.81$ $\cdot log(s) - 1.09$ $\cdot log(c) - 4.64 \cdot \rho_b$	K $[m/hours]$ c content of clay particles (in %) where $\phi \leq 2\mu$m s content of sand particles (in %) where $\phi \geq 50\mu$m	Jabro, 1992	Cronican & Gribb, 2004; 2007 Li et al., 2008 (4)

Methods	Equation	Comments	Original source	Other sources
Alyamani & Sen	$K = 1300 \cdot [I_0 + 0.025$ $\cdot (d_{50} - d_{10})]^2$	K $[m/day]$ Most suitable for well-graded sample I_0 is the intercept (in mm) of the line formed by d_{50} and d_{10} with the grain size axis	Alyamani & Sen, 1993	Batu, 1998 Odong, 2007 Li et al., 2008 Sezer, 2009 (9)
Sperry & Pierce	$K = -1.28 \cdot 10^{-1} + 9.5 \cdot 10^{-4} \cdot d_{50} +$ $7.71 \cdot 10^{-3} \cdot \gamma \cdot m$ $K = 1.1 \cdot 10^{-2} + 1.35 \cdot 19^{-6} \cdot d_{50}^2 +$ $1.14 \cdot 10^{-6} \cdot \frac{m^3}{(1-m)^2} \cdot d_{50}^2$	K $[cm/s]$ d_{50} in μm γ is the angle of inclination of the grain (in degrees)	Sperry & Pierce, 1995	Cronican & Gribb, 2004; 2007
Vereecken	$K = e^x$ $where:$ $x = 20.62 - 0.96 \cdot ln(c) - 0.66$ $\cdot ln(s) - 0.46$ $\cdot ln(om) - 8.43 \cdot \rho_b$	K $[cm/day]$ c content of clay particles (in %) where $\phi \leq 2\mu$m s content of sand particles (in %) where $\phi \geq 50\mu$m Can be applied when: $7.2\% < s < 96.9\%$ $1.1\% < u < 74.7\%$ $1.5\% < c < 51.9\%$ $0.04\% < om < 2.9\%$ $1.19\% < \rho_b < 1.73\%$	Vereecken et al., 1989	Li et al., 2008 (5)

Methods	Equation	Comments	Original source	Other sources
Kozeny	$K = \left(\frac{\gamma}{\mu}\right) \cdot \left(\frac{C \cdot m^3}{S_s^2}\right)$	Based in large grain size of sand (coarse sand)	Kozeny, 1953	Schoeller, 1962 Schwarz & Zhang, 2003 Kasenow, 2002
Barr	$K = \frac{\rho_w \cdot g}{C_c \cdot \mu \cdot C_s \cdot S_0} \cdot \left(\frac{m^2}{(1-m)^2}\right)$	K [cm/day] C_c is the shape factor (5 ± 0.25) C_s is the surface factor of the area (1.1 ± 0.1) S_0 is the area of the surface per solid mass unit * Not applied in this study	Barr, 2001	Li et al., 2008
Cronicam & Gribb	$K = e^x$ $where:$ $x = 99.386448 - 2.805685751 \cdot s$ $+0.017853198 \cdot s^2 - 0.0019859235 \cdot s \cdot c$ $+0.000268617 \cdot s^2 \cdot c$	K [cm/day] s content of sand particles (in %) where $\phi \geq 50\mu$m c content of clay particles (in %) where $\phi \leq 2\mu$m Soils with $> 70\%$ of sand	Cronican & Gribb, 2004; 2007	

Methods	Equation	Comments	Original source	Other sources
Bloemen	$K = 0.02 \cdot M_d^{1.93} \cdot GSDI^{-0.74}$	M_d average of the grain size distribution, in μm $GSDI$ is the factor of the grain size distribution. No dimensions * Not applied in this study	Bloemen, 1980	Li et al., 2008
Li	$log_{10}(K) = 5.3407 - 0.5286$ $\cdot \rho_b - 1.2846 CP$ $-0.0442 \cdot c + 0.0612$ $\cdot d_5 - 0.609$ $\cdot d_{10} + 0.085 \cdot d_{95}$ $where:$ $CP = 0.2576 - 0.0020 \cdot s + 0.0036$ $\cdot c + 0.0299 \cdot om$	K $[cm/day]$ CP is the capillary porosity s content of sand particles (in %) where $\phi \geq 50\mu$m c content of clay particles (in %) where $\phi \leq 2\mu$m	Li et al., 2008	
Bakhmeteff	$k = c \cdot d_e^2 \cdot m^n$	k intrinsic permeability $[cm^2]$ $c = 710$ $n = 4/3$	Bakhmeteff & Feodorof, 1937	Schneebeli, 1966 Custodio & Llamas, 1984

Methods	Equation	Comments	Original source	Other sources
Bakhmeteff & Feodorof	$K = \left(\frac{\gamma}{\mu}\right) \cdot \left(\frac{2 \cdot d_e^2 \cdot m^{4/3}}{c}\right)$	K $[cm/s]$ $c = 640$	Bakhmeteff & Feodorof, 1937	Schoeller, 1962 (6)
Kenney	$k = C \cdot d_5^2$	K [cm/s] C = 0.05 ∼ 1 Can be applied when: d_5 between 0.074 - 25.4 mm C_u between 1.04 - 12	Kenney et al., 1984	Sezer, 2009
Uma	$K = 6 \cdot d_{10}^2$	K $[cm/s]$ d_{10} $[cm]$	Uma et al., 1989	
Slichter	$K = c \cdot d_e^2 \cdot m^n$ $k = \left(\frac{g}{\nu}\right) \cdot n^{3.287} \cdot d_e^2$ $K = \left(\frac{g}{\nu}\right) \cdot 1 \cdot 10^{-2} \cdot n^{3.287} \cdot d_e^2$	K $[cm/s]$ k intrinsic permeability $[cm^2]$ $0.01mm < d_e < 5mm$ (From Vuković & Soro, 1992)	Slichter, 1899	Custodio & Llamas, 1984 Vuković & Soro, 1992 Kasenow, 2002 Odong, 2007 Sezer, 2009

Methods	Equation	Comments	Original source	Other sources
Breyer	$K = \left(\frac{g}{\nu}\right) \cdot 6 \cdot 10^{-4} \cdot log\left(\frac{500}{U}\right) \cdot d_{10}^2$	K [cm/s] Most suitable for poorly sorted grains Can be applied when: d_{10} between (0.06 - 0.6) mm U between (1 - 20)	Breyer, 1984	Kresic, 1998 Odong, 2007 Sezer, 2009 (7)
Chapuis(a)	$K = 1.5 \cdot d_{10}^2 \cdot \left(\frac{e^3 \cdot (1+e_{max})}{e_{max}^3 \cdot (1+e)}\right)$	K $[cm/s]$ Most suitable for clean sand and gravel d_{10} in mm Can be applied when: e_{max} between (0.7 - 0.8)	Chapuis, 2004a	Sezer, 2009
Chapuis(b)	$K = 2.4622 \cdot \left(\frac{d_{10}^2 \cdot e^3}{(1+e)^{0.7825}}\right)$	K $[cm/s]$ Most suitable for clean sand and gravel d_{10} in mm Can be applied when: e_{max} between (0.3 - 1) d_{10} between (0.003 - 3) mm	Chapuis, 2004b	Sezer, 2009 Choo et al., 2016

Methods	Equation	Comments	Original source	Other sources
NAVFAC (Naval Facilities Engineering Command)	$K = 10^{1.29 \cdot e - 0.6435} \cdot (d_{10})^{10^{0.5504 - 0.2937 \cdot e}}$	K $[cm/s]$ d_{10} $[mm]$ Most suitable for sand or mix of sand and gravel Can be applied when: C_u between (2 - 12) d_{10} between (0.1 - 2) mm e between (0.3 - 0.7) mm	DM7, 1974	Chapuis et al., 1989b Chapuis, 2004a sezer, 2009 Chapuis, 2012
Choo	$K = \frac{\gamma_w}{5 \cdot \mu} \cdot \left(\frac{d_{eq}}{SF}\right)^2 \cdot \frac{F^{-3/m}}{(1 - F^{-1/m})^2}$ $where:$ $d_{eq} = \frac{100\%}{\sum(f_i/(d_{Li}^{0.404} \cdot d_{Si}^{0.595}))}$ $F = \frac{\rho_{mix}}{\rho_w = a \cdot S^{-t} \cdot m^{-v}}$	K $[cm/s]$ F is the formation factor * Not applied in this study	Choo et al., 2016	
Rosetta	ANN		Schaap et al., 1998; 2001	

Methods	Equation	Comments	Original source	Other sources
USBR	$K = \frac{g}{\nu} \cdot C \cdot d_{10}^2$ $where:$ $C = 4.8 \cdot 10^{-4} \cdot d_{20}^{0.3}$	K $[cm/s]$ d_{20} $[mm]$ Most suitable for medium-grain sand Can be applied when: $C_u < 5$	No references could be found. Some authors cited it after Malet & Pacquant (1954)	Vuković & Soro, 1992 Odong, 2007 Cheng & Chen, 2007 Sezer, 2009 (8) Vienken & Dietrich, 2011
Shahabi	$K = 1.2 \cdot C_u^{0.735} \cdot d_{10}^{0.89} \cdot \frac{e^3}{1+e}$	K $[cm/s]$ d_{10} $[mm]$ Most suitable for sand Can be applied when: $1.2 \leq C_u \leq 8$ $0.15 \leq d_{10} \leq 0.59$ mm $0.38 \leq e \leq 0.73$	Shahabi, 1984	Chapuis, 2012
Mbonimpa	$K = C_G \cdot \frac{\gamma_w}{\mu_w} \cdot$ $C_u^{1/3} \cdot d_{10}^2 \cdot \frac{e^{3+x}}{1+e}$	K $[cm/s]$ d_{10} $[cm]$ $C_G = 0.1$ $x = 2$	Mbonimpa, 2002	Chapuis, 2012

Notes

(1) In the absence of the original source: maybe the parameter u is per unit, instead of percentage. With that propose the error was reduced some orders of magnitude.

(2) There is a numerical mistake in that paper that was detected comparing with Saxton (1986); the coefficient 0.06671 is in fact 0.03671. In this study it was corrected.

(3) There is not any specification about the application domain in the original source.

(4) There is a coefficient in the equation that Li et al (2009) defined as sand percentage (s) and Cronican and Gribb (2007) as silt percentage (u). In the absence of the original source, this study used the Li et al. (2009) definition because it worked better.

(5) There are two mistakes in that paper, that was detected comparing with Vereecken (1989); the brackets are wrongly situated and the reference of source mistakes the publish year. In this study all were corrected.

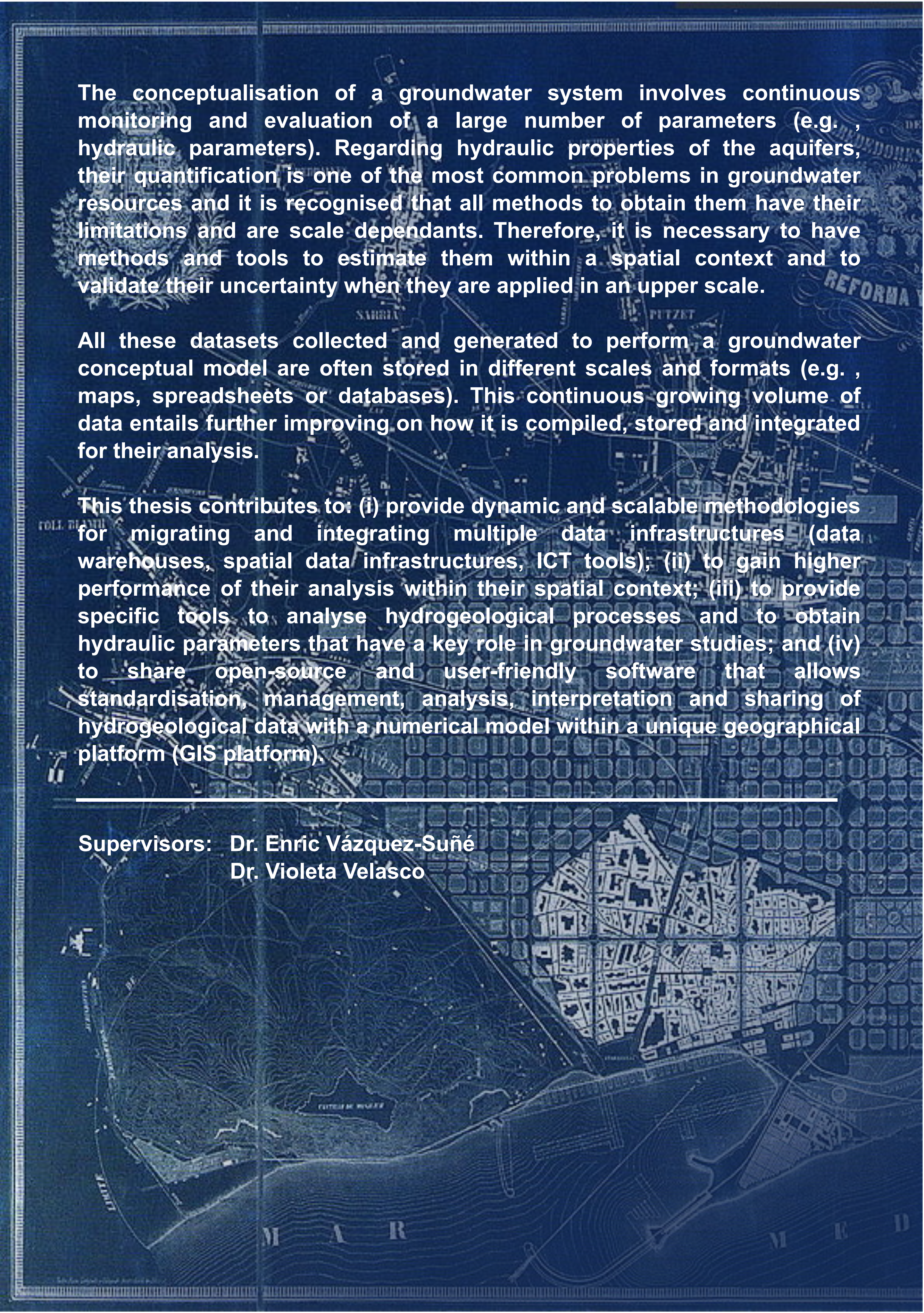
(6) In this reference the hydraulic conductivity is defined in (gallon/day). It make no sense because that corresponds to cabal units and it must defined in velocity units. In that study it was converted to (cm/s) and the results were satisfactorily.

(7) In this reference (Sezer, 2009) the constant $6 \cdot 10^{-2}$ is two orders lower than Odong, 2007 ($6 \cdot 10^{-4}$). In that study it was used $6 \cdot 10^{-2}$ and the results were satisfactorily.

(8) In this reference (Sezer, 2009) the constant ($4.8 \cdot 10^{-2}$) is two orders lower than Vukovic & Soro, 1992 ($4.8 \cdot 10^{-4}$). In that study it was used the original formula collected from Vukovic & Soro, 1992. The results were satisfactorily.

(9) In this reference (Sezer, 2009) the constant (1.5046) is three orders lower than Odong, 2007 (1300). In that study it was used the original formula collected from Odong, 2007. The results were satisfactorily.

(10) Vukovic & Soro, 1992 suggest K in $[cm/s]$ with ν (viscosity) in $[m/s]$. Kansenow, 2002 suggest K in $[m/s]$ with ν (viscosity) in $[m/s]$. We used K units proposed by Kasenow for coherence in the units.



The conceptualisation of a groundwater system involves continuous monitoring and evaluation of a large number of parameters (e.g. , hydraulic parameters). Regarding hydraulic properties of the aquifers, their quantification is one of the most common problems in groundwater resources and it is recognised that all methods to obtain them have their limitations and are scale dependants. Therefore, it is necessary to have methods and tools to estimate them within a spatial context and to validate their uncertainty when they are applied in an upper scale.

All these datasets collected and generated to perform a groundwater conceptual model are often stored in different scales and formats (e.g. , maps, spreadsheets or databases). This continuous growing volume of data entails further improving on how it is compiled, stored and integrated for their analysis.

This thesis contributes to: (i) provide dynamic and scalable methodologies for migrating and integrating multiple data infrastructures (data warehouses, spatial data infrastructures, ICT tools); (ii) to gain higher performance of their analysis within their spatial context; (iii) to provide specific tools to analyse hydrogeological processes and to obtain hydraulic parameters that have a key role in groundwater studies; and (iv) to share open-source and user-friendly software that allows standardisation, management, analysis, interpretation and sharing of hydrogeological data with a numerical model within a unique geographical platform (GIS platform).

Supervisors: Dr. Enric Vázquez-Suñé
Dr. Violeta Velasco